Eötvös Loránd University, Budapest
Biology Doctorate School, Head: Anna Erdei, DSc
Ethology Doctorate Program, Head: Ádám Miklósi, DSc

Dog *(Canis familiaris)* growls as communicative signals

Acoustic and behavioural analyses

PhD. Thesis

Tamás Faragó

Supervisor: Péter Pongrácz, PhD

ELTE Department of Ethology
1117 Budapest, Pázmány Péter sétány 1/c

2011

Contents

“The single biggest problem in communication is the illusion that it has taken place.”

George Bernard Shaw

Introduction

Vocal communication

Defining communication

During their lives, every single animal gets into different interactions with other animals. In these interactions specific behaviours, markings and other attributes act as an agent between the animals. In ethology communication is assumed if one individual (the initiator) modifies the behaviour of another (receiver) by using specially evolved signals or displays (Dawkins & Krebs 1978). The initiator is usually called as actor or signaller (sender) because it plays the role of the starting point of the communication process; it is the source of the signal that can affect the behaviour of the other individual called the receiver or reactor that responds to this signal (Figure 1). Easy to imagine a wolf (*Canis lupus*) approaching a feeding dominant conspecific: as the dominant one will start to growl and show its bare teeth, the first one soon shows submissive behaviour and retreats. So the receiver was able to decode the agonistic signals of the actor, and due to it, changed its way of behaviour. Finally, both participants can potentially benefit from the interaction because the receiver might be able to predict the sender's future behaviour and modify its own behaviour according to this information, on the other hand the sender might have the opportunity to alter the receivers behaviour and gain from the interaction (Bradbury & Vehrencamp 1998).

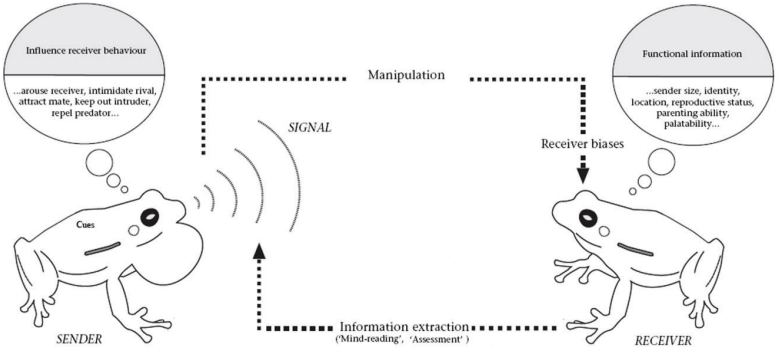

Figure 1. The schematic overview of communication (Modified graph of Font & Carazo 2010)

In the example above it is clear that the behaviour of the dominant individual deterred the other, but other features such as body size or the length of the canine teeth per se can have a deterrent effect, and this way can also change the behaviour of the other party. However these latter features did not evolve for this 'purpose' mainly, thus are not mentioned as signals, but cues (Goodenough et al. 2001a). There are major differences between cues and signals. While signals are not always present and the individual have some control on showing them, cues cannot be ‘switched off’, they are phenotypic and provide information about the individual without additional energy investment (Hauser 1997). One other difference between cues and signals that in most cases cues cannot be behaviourally modified (although, we will mention an example of cue modification in the later chapters) (Donath 2007).

Signals are also arbitrary, in the meaning that the signal form has no obvious relation with the meaning it conveys (Barnard 2004). Although the roaring of a red deer (*Cervus elaphus*) do not delineate the harem with its ruler, several cues can be encoded within the sound which were formed by evolutionary constrains and convey information about the signaller such as physical abilities, size, strength, level of excitement (however in some cases the signal can foreshadow an event that is in strong relation with the signal, such as bare teeth display carries the chance of the use of the teeth). Due similar constrains, signals also may be linked and closely related to their producers. Signals emerged from other behaviours such as intentional movements through a process called ritualization. During this process, when a movement acquire information-provider functionality, it will be exaggerated, become repetitive and highly stereotyped through evolution. The selective pressure for its distinction from the original behaviour will push this ritualized behaviour to become a communicative signal (Hauser 1997). Evolutionary constrains caused by the anatomical structures of the signaller, or originating from the environment can enforce the emergence of signals with universal traits, like alarm calls that must be hard to localize, but highly audible, therefore these calls are high pitched and short in several far related species. Also the external world can have a strong impact on signals, therefore strong selective pressures formed them to carry information as reliably as possible. For example, torrent frogs (*Amolops tormotus*) living in fast running, cataract rich streams have to cope with the constant, wide frequency ambient noise, thus the males produce their advertisement calls in ultrasound range to overcome the noise, and also their hearing is tuned to these ultrasonic frequencies (Arch et al. 2009).

Graded and discrete signals

We can divide the signals into two subgroups depending on their way conveying information: discrete or graded signals (Sebeok 2007). Graded signals provide information by their differences in intensity and complexity (Barnard 2004). Mostly the intensity pattern is in strong relationship with the inner state of the signaller and the evoking stimulus. Good examples are the aggressive displays of several social mammalian species, for example the facial expressions of the cats (*Felis catus*) can be arranged along an aggressiveness-fearfulness scale (Leyhausen 1973), also in the elephants (*Loxodonta africana*) the position of the trunk and the ears predict reliably the likelihood of attack or flight (Kühme 1963). Thus graded signals can be sequential as appearance of different behavioural elements shows the progress or the elevation of the inner state. Other possibility when the inner state of the signaller is projected by physical characteristics of the signal. Some acoustic parameters of a vocalization can correlate with the inner state of the signaller: for example the lower the pitch of the sound, the higher the aggression of the signaller (Morton 1977).

The other type of signals called discrete signals are considered to work in a digital, on/off manner, or operate with some few distinct signal types. The light flashing signalling of the firefly females that attract the males is a reasonable example, because not the length of the flashes, or the intensity is that provide information, just the rhythm of the sequences (Barnard 2004). Different firefly species use similar light spectra for signalling, and they not able to modify the intensity of the light emittance, they just control the on/off state of the light signal. Thus the only possible way for species to differ is to select for different flashing sequences among the species. This way the neighbouring species can avoid interference and mating with an extra-species partner. Such highly stereotyped behaviours can be considered as discrete signals, because they are characterised by a typical intensity independently from the current inner state of the signaller (Morris 1956). The alarm call systems of different species rely also on discrete signals because each predator specific alarm call works as a separate signal.

Acoustics

The vocal communication uses the auditory channel to convey information by sound waves. Sound waves are “the propagation of density changes through an elastic medium [or] the transfer of energy through an elastic medium” (Speaks 1992). The medium can be any kind of material that has an ability to vibrate, which, in case of animal communication

usually includes air, water or the ground. Thus sound waves are longitudinal vibrations that travel through the medium from the source (the signaller's vocal apparatus) to the receiver's auditory system. During the travel mechanical energy is delivered by cyclic collisions of the molecules of the medium. These can be imagined as attenuations and condensations of the particles, but it is important to emphasise that not the particles travel just the energy during this transmission (Figure 2). These pressure changes can be characterized in both frequency and time domain as sinusoid resonances. The frequency (f) is the number of cycles of the condensations in every second measured in Hertz (Hz). The reciprocal of this frequency (1/f) called period (T), is the duration of one cycle, while the distance covered by one cycle is the wavelength (λ) defined in metres and depends on the speed of sound propagation (c) in the medium ($\lambda = c/f = cT$) (Simmons et al. 2003). Natural sounds are usually not clear sinusoids like tones, but complex waves that can be considered as superimpositions of different frequency and amplitude sinusoids which can be revealed by a mathematical process called the Fourier transformation (Bradbury & Vehrencamp 1998).

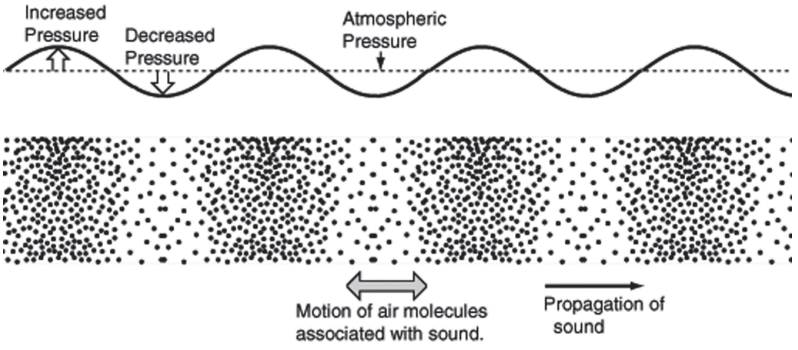

Figure 2. The propagation of the longitudinal sound waves in gas. The cyclical pressure changes travel through space, while the gas particles remains in a relatively constant space (from: http://hyperphysics.phy-astr.gsu.edu/hbase/sound/tralon.html)

When we pluck a string on a guitar, thanks to its elasticity it will move up and down periodically in its full length on a given frequency, which depends on its thickness, length and tension. This frequency is called the fundamental frequency (F0), because it gives the base of the sound. While only the two endpoints are fixed, the string will not just move up and down in whole length, but certain points of it can remain still, while the two opposite side will move alternately (Figure 3). These resting points are at the integer deviates of the

string length, and these additional vibrations will give integer multiples of the fundamental frequency called the harmonics of the sound.

When we record the sound of the vibrating string and graph it by time and amplitude dimensions we will see a complex waveform called oscillogram (Figure 4a). During its analysis we can dissolve this wave with Fast Fourier Transformation to its sinusoid frequency components and we will see on which power they contribute in the composition of the sound wave. When we graph the time frequency and power dimensions, we will get the spectrogram or the two dimensional version called sonogram of the sound (Figure 4b). On this graph the lowermost and strongest energy frequency part will present the fundamental frequency, and the harmonics will appear as equally distributed collateral lines above the fundamental. These two graphical representations besides various computer technics can help a lot in the analysis of sound waves. Obviously natural sounds can be far more complex than the sound of a string, because they contain a lot of noise, irregular changes and nonlinear elements that add more "shades" to the spectrum.

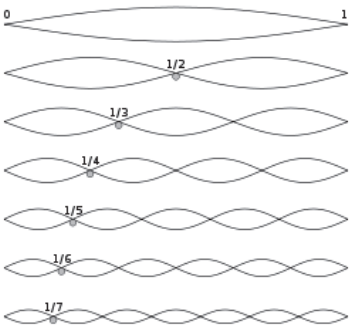

Figure 3. Vibrations of a string. The resting points divides the length of the string with consecutive integers creating the upper harmonic vibrations on multiplied frequencies of the resonance of the full length string

(From:http://upload.wikimedia.org/wikipedia/commons/c/c5/Harmonic_partials_on_strings.svg)

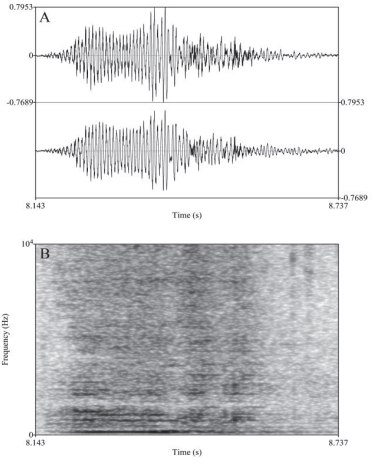

Figure 4. A, The oscillogram of a dog growl. B, the sonogram of the same sound. The oscillogram shows the actual waveform of the signal, while on the sonogram the frequency spectrum can be studied (the author's recording, graph created with Praat).

Sound production

A vast number of species in animal kingdom use the acoustic domain for communication, but it is mostly typical among particular arthropods and vertebrates.

Acoustic signals can travel large distances in air, water and ground in all directions, and carry various amount and type of information between the signaller and receiver, and also can have impact on the behaviour of third party conspecifics and individuals from different species such as predators (Simmons et al. 2003). Moreover, the characteristics and changes of the environment can distort the acoustic signal, add or mask it with noise (Barnard 2004). These environmental features had an important role in the evolution of the acoustic communication.

The Source-Filter Theory of sound production

The Source-Filter Theory attempts to describe the mechanism of sound production based on the studies of human speech, and it provides a good framework for studying vocal communication (Taylor & Reby 2010). Speech scientists revealed that the human voice production is a two stage process, and in these two part, different elements of the vocal apparatus are involved: the source and the filter (in details see above), both acts as two discrete but not independent functional unit during voice production (Fant 1960; Titze 1994). While the basic anatomy of the vocal apparatus in terrestrial vertebrates is fundamentally same, the theory could be generalized to cover the sound production mechanism of nonhuman species either (Fitch 1994).

Energy

The basic principles of actively creating sounds are the same among invertebrate as well as vertebrate species (Fitch & Hauser 2003). The mechanical energy evoked by muscle movements is transformed by some flexible membrane or fold by its vibration to sound energy that can be transmitted by the cyclic pressure changes of the medium to finally reach the recipients' auditory organs. This happens when a cricket frays together particular edges of the wing for creating that typical chirping sound in order to attract females, when a frog pumps back and forth the air between its air sac and lungs to get its arytenoid discs to resonate in the larynx, when a songbird pushing air out of its lungs to flow through the two pairs of syringeal membranes generating a whistling sound, and also when a deer breathe out forcedly and the vocal folds starts to open and clap together repeatedly in its larynx while roaring to scare away its competitors. In terrestrial vertebrates the source energy originates from the muscle contractions causing the rise of the air pressure in the lungs (Liem 1985). This pressure will cause the air flowing out through the bronchia into the main source of the acoustic signal. In most vertebrates this source is the larynx, while

in birds the syrinx, an evolutionary novel organ specialized for vocalization plays a similar role, (Simmons et al. 2003). Additionally air sacs linked to the respiratory system can function as air reservoirs, saving energy of the respiration and they can have other important acoustic functions too (Riede et al. 2008). For example, in anuran species besides this recycling of air used for vocalization the air sacs function as impedance-matching device and increase the efficiency of transferring the acoustic energy to the environment (Narins et al. 2007).

Source

The main sources of the vocal signals are elastic, mobile structures that can be actively placed in a position to the way of the airflow, and are able to vibrate. In birds, these structures are the tympaniform membranes or the syringeal labia that take place on the medial and lateral wall of the two bronchi just under the exit of the trachea, while in other tetrapods the vocal folds in the larynx are the mechanical vibrators (Fant 1960; Titze 1994; Truax 2001; Mindlin & Laje 2005). During voicing, the two folds or membranes come close together as a result of the specific muscle movements and they block the airflow from the lungs. This will cause the rise of the pressure under the closed vocal folds till the air will be able to push away the folds again, and it will starts to flow out between them. Thus, the air flow is accelerated in the opening, which causes a local decrease of pressure that sucks together the vibrators again, and together with the elastic movements it will make the folds collide. This collision of the two vibrators generates the acoustic energy (Figure 5). After this the continuous flow of air will raise the pressure under the vibrators again that cause the opening of the two

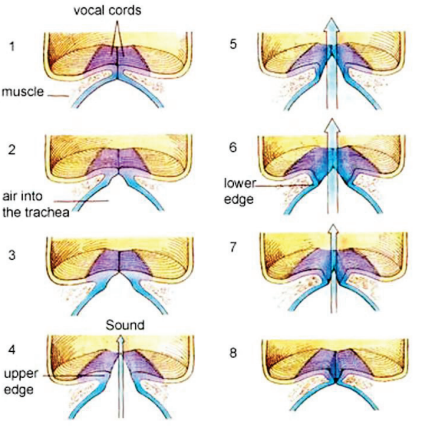

Figure 5. The cross section of the vocal folds (or cords) during a cycle of sound production. 1, the folds are closed and blocking the flow of air from the lungs. 2-3, the raising pressure starts to push the folds away from its way. 4-6, the opening between the two folds lets out a puff of air. 6-7, the folds start to collapse due to their flexibility, inertia and the lowering air pressure in the flowing air between them. 8, the folds closed together again

(From:http://www.singintune.org/voice-production.html)

vocal folds and the cycle starts over. Important to consider that the closing and opening of the vibrating parts is a passive process caused by the pressure changes in the airflow and the inertia and the elasticity of the folds. The muscle work is only necessary for the positioning of vibrating parts and the generation of the airflow from the lungs.

Thus, the cyclic clapping of the vibrators generates the basis of the acoustic signal. The rate of this pulse gives the fundamental frequency of the sound. If we record only the sound of this source without the upper tracts of the respiratory system, we will get only a humming sound containing the frequency band that have the most energy, the fundamental frequency and its upper harmonics (Figure 6). The rate of the fundamental frequency is mainly determined by the mass and elasticity of the vibrators, the pressure in the lungs, and the tension state of muscles involved in positioning the vibrators (Fitch & Hauser 2003). Modifying the state of these muscles can be a way for the animal to adjust the fundamental frequency of the vocalization, although their repetitive constriction is rarely used for sound production because the muscles are not able to twitch at so high frequencies without going into tetany. One exception is the purring of cats that has the fundamental frequency at 25Hz and is generated by active muscle movements in the larynx (Remmers & Gautier 1972).

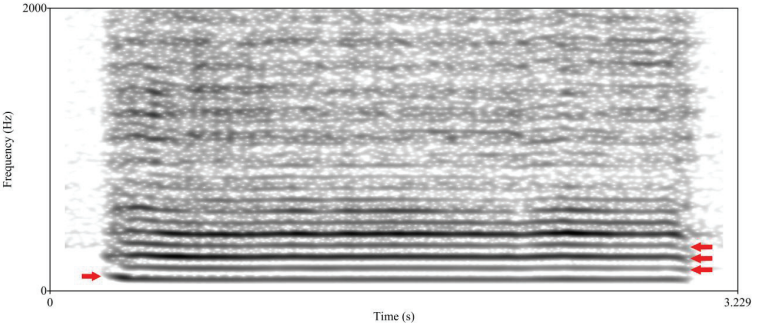

Figure 6. The spectrum of a relatively tonal sound generated by blowing air into a long tube (the author's recording). The left red arrow shows the lowest frequency component, the fundamental frequency at 100 Hz. The right red arrows show the upper harmonics at 2-3-400 Hz.

As a complex natural system, the oscillations of the source are never regular and contain chaotic, nonlinear elements (Fitch & Hauser 2003). Besides the pathological cases, these normal irregularities can have a communicative role too (Wilden et al. 1998). In the last two decades studies on occurrences of nonlinear phenomena in animal vocalizations shed light on this role (Fitch et al. 2002). These phenomena, due to their nonlinear nature

are unpredictable, and can highly enhance a vocal signal's attention grabbing power. These nonlinear events can make it hard to become habituated to these calls (Figure 7). Therefore calls emitted in frustration, alarm and begging calls usually contain such elements. These nonlinearities can also act as cues for individual recognition because they are highly dependent on the individual variability of the anatomical structure of the vocal apparatus. Due to this they can also carry information about the individuals' quality: fluctuating asymmetries that appears during the development of the individual can enhance the appearance of chaotic events because asymmetric vocal folds tend to vibrate in asynchronous way causing biphonation or other irregularities. On hearing a signal containing nonlinear elements, these chaotic changes will be processed as subjective impression of changes in noisiness. This can be measured with the Harmonics-to-Noise ratio (HNR) of the signal, that is defined as the tonality: low HNR indicates a harsh, atonal sound while high HNR tonal, clear sounds (Yumoto et al. 1982).

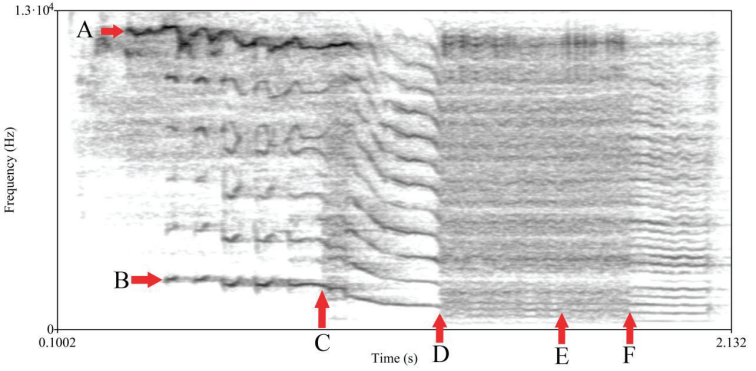

Figure 7. Occurrences of nonlinear phenomena in a whine of a dog during separation from owner (the author's recording). A, Biphonation, a secondary fundamental frequency produced in the nasal tract of the dog. B, the main fundamental frequency with strong frequency modulation. C, Frequency jump with additional noise component. D, Noisy subharmonics. E, Deterministic chaos caused by the irregular resonances of the vocal folds. F, tonal subharmonics called frequency doubling caused by the asynchrony of the two vocal folds.

Other ways of sound production are also possible like constricting the vocal tract at specific places to cause perturbations in the airflow, which can cause hissing or whistling sounds, or using different parts of the body like the tail clashing of beavers (*Castor sp.*) as alarm signal, or the threatening chest beating of the male gorillas (*Gorilla sp.*) (Fitch & Hauser 2003); however in the further discussion we will concentrate on the more conventional form of vocalizations of terrestrial vertebrates.

Filter

Once the acoustic energy was generated in the source, the vibrations of air to reach the environment have to pass through the upper parts of the respiratory system, called the vocal tract. The vocal tract in birds includes the trachea, the larynx and the oral cavity, while in other tetrapods only the supra-laryngeal parts of the respiratory system and optionally air sacs in addition serve as vocal tract. The vocal tract also shapes, modifies the passing acoustic signal (Fant 1960). The air column inside the vocal tract preferentially vibrates at certain frequencies, called normal modes or resonances, due to physical constrains of its given elasticity and mass (Figure 8). These vibrations interfere with the source signal, they enhance given frequency bands of it while fade out others, thus the vocal tract acts like a frequency band filter. In human speech, these enhanced frequency bands, called formants are responsible for the shaping of different vowels (Fant 1960). The arrangement and shape of the formants depends only on the shape and length of the vocal tract, thus their modification can carried out by changing the length of the vocal tube directly, or indirectly with changing the openness of the lips, or closing the mouth and using the nasal cavity, and also by modifying the shape of the tube with positioning the lips and the tongue (Fitch 1999; Fitch 2000; Fitch & Reby 2001).

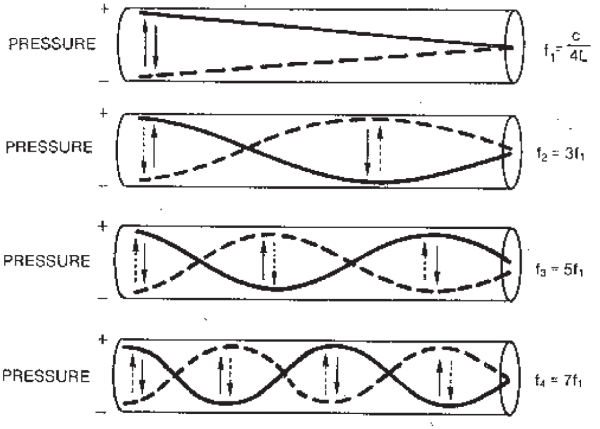

Figure 8. Resonance frequencies of a tube closed in one end. Waves with given length can effectively travel through the tube. Their frequency depends on the length of the tube. The enhancement of these frequencies and the attenuation of the not efficient ones produce the formant frequency bands in the signal. (http://home.cc.umanitoba.ca/~krussll/138/sec4/enhance.htm)

One important feature of this basic model of sound production is that the source and the filter work independently from each other. This means that modifying the fundamental frequency will not affect the shape and placement of formant frequencies, and vice versa (Fitch & Hauser 2003). This can be easily understood if we think about human speech (Figure 9). One is able to produce the same vowel while changing the tone of the voice, and also can form different vowels with a fixed frequency. Further evidence is offered by sound production experiments in different species in which subjects breathe light or heavy gases. In heliox environment for example, the fundamental frequency of the vocalization will not be affected, while the formant frequencies will slide higher in the spectrum (Beil 1962; Rand & Dudley 1993). This independence of the source and filter supports high variability in vocalization of animals and humans.

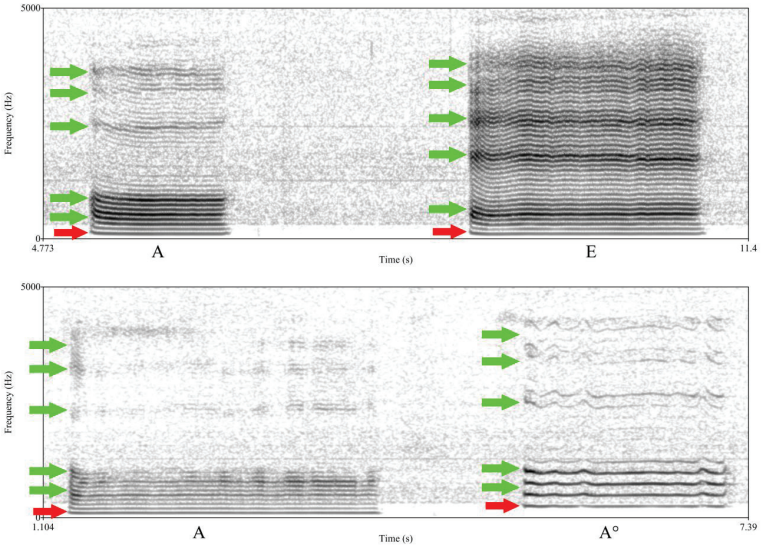

Figure 9. The spectrum of two vowels (A and E) showing the independence of source and filter (the author's recording). Red arrows point at the fundamental frequency, the green arrows at the formant frequencies. On the top picture the fundamental frequency is constant, while the change of the second formant differentiate the two vowels. On the bottom picture, the formants are unchanged, while the fundamental frequency rose from 100 Hz to approximately 200 Hz.

The signal

In two basic domains can an acoustic signal convey information: in frequency and time (Fitch & Hauser 2003). The frequency structure of the signal depends on the composition of the sound source and the vocal tract: it can be modified by several muscular movements

during sound production. On the first level, the source signal can be modulated by positioning and varying the tenseness of the vocal folds in the larynx. The closer the folds are pushed together, the thicker the folds are and the more mass have to be moved by the floating air, the lower is the frequency of sound produced. The vocal folds have no intrinsic musculature; they can be moved or modified by the movements of cartilages in the larynx.

Modifying the filter function of the vocal tract provides also a suitable way of forming the spectral appearance, and this way the sounding of a vocal signal. There are two main possible ways, the active and passive modifications. The latter is happening in evolutionary time caused possibly by some selective pressures in mate choice or/and conflicts. There is an elongated vocal tract in several avian species like trumpeter swan (*Cygnus buccinator*), crested guinea fowl (*Guttera edouardii*) while in several mammalian and other species the lengthening of the supralaryngeal parts with proboscis can be found (Fitch 1999; Fitch 2000; Fitch & Reby 2001) (Figure 10a). The active modification can be the elongation of the vocal tract by muscles pulling down the larynx during vocalizations (Figure 10b), or the shaping of the vocal tract, by movements of the hypoglottis, the tongue, the lips, opening or closing the mouth, or detour the airflow from the oral to the nasal cavity (Hauser et al. 1993; Reby et al. 2005). All these mechanisms affect the formant shape and position in the vocalization affecting the sound itself. Besides these the amplitude, the strength of the signal also can carry information, however these depend mainly on the respiratory capacity and the amount and strength of muscles involved in sound production.

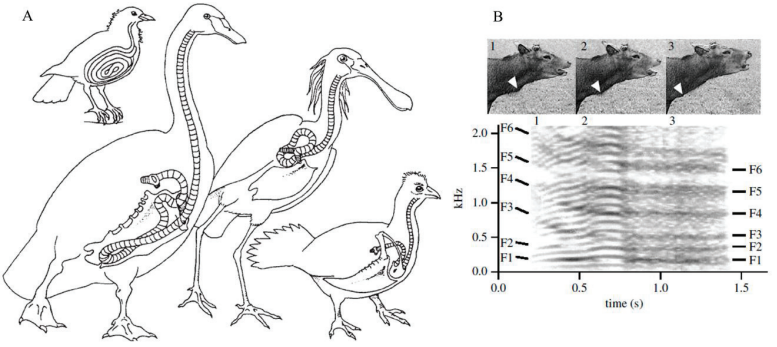

Figure 10. A, Elongation of the vocal tract in several different bird species (from left to right: trumpet manucode (*Manucodia keraudrenii*), trumpeter swan (*Cygnus buccinator*), European spoonbill (*Platalea leucorodia*) and crested guinea fowl (*Guttera edouardii*)) (from Fitch 1999) B, Active elongation of the vocal tract in red deer, by the retraction of the larynx (from Fitch & Reby 2001).

Information content in vocalizations

Vocalizations can provide cues about the caller because of the physical constrains during their production, about the motivational state of the signaller caused by the neural control modified by emotional states, and convey information in an arbitrary manner such as about territory ownership, identical and species linked cues, and external references such as predators and food sources.

Vocal cues of the caller's attributions

During several social interactions like agonistic encounters and mate choice, it is important to broadcast and gain information about the physical parameters, age or dominance status of others (Owings & Morton 1998). In both agonistic contexts and mating a two-way selection pressure is working on the vocal communication system that was often portrayed as an arms race (Slatkin & Smith 1979). Although the physical constrains limit the possibility of dishonest signalling, the signallers may evolve ways to provide fake information about themselves. In parallel the receivers can modify their assessment to counteract dishonest signalling (Fitch & Hauser 2003). This situation can be caught in the act in the roars of the red deer. David Reby and co-workers provided several experimental evidence which support the arms race theory (Reby & McComb 2003a). First measuring acoustical parameters of the roars of the red deer stags during the mating season showed that these vocalizations carry honest information of the body size which is a reliable cue to assess the fighting ability of the animal (Reby & McComb 2003b). The formant dispersion of the roars is in a strong relationship with the length of the vocal tract during vocalization; lower formants and dispersion correlates with longer vocal tract and larger body size. While the vocal folds that are the source of the call grow and vary in mass and length independently from other parts of the body, the vocal tract length is dependent on the skeletal structure of the skull and neck, thus limited by body size. This limitation ensures that the format dispersion is a honest vocal indexical cue of the body size (Reby & McComb 2003b). This effect is not restricted to this species but has been revealed in several mammals and birds (e.g. dogs, rhesus monkeys, whooping cranes and humans), and theoretically it may be a general feature for terrestrial vertebrates (Fitch & Reby 2001).

It was also showed that red deer hinds are sensitive to the formant structure of the male's roars (Charlton et al. 2007b). Deer hinds in oestrus showed a clear preference towards played back roars that had lower formants and dispersion in a two-way choice test (Charlton et al. 2007a). Moreover, this preference was unchanged when the playbacks

indicated artificially enlarged males that have physical parameters out of the natural range. This finding suggests that the females have an ‘open-ended’ preference towards larger males, what is a typical phenomenon that emerges during a 'run-away' process caused by sexual selection (Fitch & Hauser 2003). Here the larger body size correlates with better genetic quality, thus the choice of the bigger male gives a selective advantage because it results in better quality calves (Clutton-Brock et al. 1979). This preference of the females put a selection pressure on the males advertising behaviour, and in some point of evolution the ancestor of the red deer achieved an ability to exaggerate their apparent body size in their vocalizations. This exaggeration is caused by the descent of the larynx to a deeper position in the neck than it is usual in nonhuman mammals. Thus deer have an unexpectedly long vocal tract for their body size. Furthermore, they are able to lengthen actively their vocal tract by pulling down their larynx from the resting position to the sternum during vocalization with a specialized muscle, this way they are able to indicate a larger body size than the real one (Reby & McComb 2003a). However, roaring per se is an exhausting activity, but retracting the larynx put in further energy cost, thus only the individuals with better physical abilities can really exploit this ability, and also smaller individuals could lengthen their vocal tract only in a shorter range than larger ones, thus larger and better quality males still have their advantage (Reby & McComb 2003a).

It is well known that the size of the body is in correlation with the success in agonistic encounters and also in reproduction (e.g. red deer: (Clutton-Brock et al. 1979), elephant seals (*Mirounga angustriostris*): (Cox 1981) and cichlid fish (*Cichlidae*): (Enquist et al. 1990)). Thus the decision to fight or flight depends on the others relative size compared to one’s own. The assessment of the body size of the other based on vocal cues, and trying to indicate a larger body size show ourselves bigger than the opponent are both evolutionary adaptive strategies. Free ranging red deer stags show more attention toward those vocalizations indication larger individuals, and also vocalize back with higher rate (Reby et al. 2005). Reby and co-workers found that these reply vocalizations are also different, depending on the apparent size of the modelled stag. Vocalizations indicating bigger body size evoked noisier, harsher roars from the subjects that can be linked with stronger motivational affect. Moreover, when they heard roars indicating larger individuals, they pulled back their larynx more, trying to show themselves even bigger in their calls.

In red deer, the mate choice also put an additional selection pressure on the males’ vocal behaviour. Anatomical constrains (the position of sternum at the bottom of the neck

of the deer stags blocks the further retraction of the larynx) limit the possibility for displaying size exaggeration. Thus individuals' indication of larger size has an upper limit which provides the opportunity for the receivers to assess the real size of the callers. This suggests that the roars of the red deer are still honest indexical cues (Reby & McComb 2003b). Since this extensive work on the red deer, similar anatomical and behavioural phenomena have been reported in several other polygynous ruminants like fallow deer (*Dama dama*) (McElligott et al. 2006), Mongolian gazelle (*Procapra gutturosa*) (Frey & Gebler 2003), saiga (*Saiga tatarica*) (Frey et al. 2007) and the goitre gazelle (*Gazella subguttorosa*) (Frey et al. 2011).

Individual recognition

For animals living in large, closed social groups it has a high importance that they are able to recognize individuals based on their vocalizations (Trivers 1971). Individual recognition may be also crucial in species who leave their kin alone and hidden for longer periods when they go collecting food, because during the reunion they should be able to recognize each other before visual contact (Trivers 1971). In these species we usually find specific vocalizations for identification of identity. Several acoustic features of the calls may be characteristic to the individual. These are mostly different distortions which affect the signal during voice production which relate to the characteristic differences of the individuals' vocal tract (Taylor & Reby 2010). These effects can be the described by the noise profile of the vocalization, the presence of different nonlinear events (e.g. deterministic chaos or subharmonics), and include temporal features (e.g. rhythm, tempo and duration) which depend on the lung capacity and the musculature of the vocalizing individual (rhesus (*Macaca mulatta*): (Rendall et al. 1996); dogs: (Yin & McCowan 2004; Molnár et al. 2008); fur seals (*Callorhinus ursinus*): (Charrier et al. 2003); roe deer (*Capreolus capreolus*): (Reby et al. 1998)). Still, the most common feature determining the individuality of the calls is the active modification of the spectral structure, especially the fundamental frequency (baboon (*Papio cynocephalus ursinus*): (Owren et al. 1997); bottlenose dolphin (*Tursiops truncatus*): (Janik et al. 2006); fallow deer: (Torriani et al. 2006); wolves: (Palacios et al. 2007)). These individual specific frequency modulations ("vocal signatures") can be used for individual recognition within the group, and for recognise non-group members. As group members are able to recognize each other's signatures, when they hear a species specific call with unknown fundamental frequency

structure, they can identify its caller as a stranger, and handle it as an intruder in the territory (Holekamp et al. 1999).

Vocal signals play an important role in kin recognition which may take place in different contexts. For example, in some species the mothers hide their offspring in the bushes during foraging, and it is advantageous for the young to remain silent until she returns. In some cases the offspring may change the hiding location in the absence of the mother. In this situation it is advantageous for the mother to signal their return, and the youngster to recognise the specific call of their mother (Torriani et al. 2006). The fallow deer provides such an example. The fawns remain hidden till they hear their mother's call that contains a specific fundamental frequency pattern. However, in this case the individual recognition is asymmetric because the mothers are not able to recognize their fawns' vocalizations. In other species in which the offspring stays in a group of other similar aged youngsters ("kindergartens"), the mutual recognition is essential (fur seals: (Insley 2001); banded mongoose (*Mungos mungo*): (Müller & Manser 2008); sheep (*Ovis aries*): (Searby & Jouventin 2003)). There are two different breeding strategies in penguins. Some species use nests during breeding for example the Adélie penguins (*Pygoscelis adeliae*), while the others are non-nesting (e.g. the emperor (*Aptenodytes forsteri*) and the king penguins (*A. patagonicus*)) in which the chicks remain in kindergartens while the parents are collecting food. In the nesting species Aubin and Jouventin found that the chicks recognized their parents based on the fundamental frequency structure of the calls. In contrast, the non-nesting species use complex coding based on the combination of frequency and amplitude modulation that result in a more sophisticated system of individual recognition (Lengagne et al. 2001; Jouventin & Aubin 2002). This difference is probably related to the different breeding ecology of the two groups of species. In the nesting species the place of the nest provides spatial cues for the parents for finding the chicks, but the non-nesting species have to rely on mutual individual vocal recognition of chicks and parents (Jouventin & Aubin 2002). Species breeding in colonies have to battle with the ambient noise and the noise caused by the superposition of the neighbours' calls. Thus these species show surprisingly efficient skill to recognize their kin's vocal signatures in the loud noise of others (Aubin & Jouventin 1998).

The spectral structure of calls provides other possible ways of individual recognition. The formant frequency differences are used in several mammalian species including red deer, rhesus monkeys and also humans (Reby et al. 2001). This effect is caused mainly by the morphological differences in the structure of the vocal tract that is responsible for the

shaping of the formant frequencies. These format patterns provide the possibility of individual recognition from not just specific calls, but having a distinctive, general voice independently from the call structure. Due to the independence of the filter in call production the formants can remain unchanged in calls with different fundamental frequency structure (Taylor & Reby 2010). This is evident in humans because we are able to discriminate speakers independently from the sounds or words produced by the person.

Communicating inner state

During social interactions, animal vocalizations are often related to the emotional or inner state of the animal (Ohala 1996). In these interactions, maybe inner state communication can be more relevant, as most of the time the participants can have visual information about each other to assess physical abilities. In such cases the acoustical cues can be redundant or overwhelmed by visual ones. The vocal signals used in different contexts can vary within and call types. This variance is clearly a result of the modification of the acoustic structure of these calls. For example within dog barks there is a high variance in several acoustic parameters like the mean, fundamental and peak frequencies, the tonality, the rhythm, and along these parameters the dog barks can be scaled by the emotional state (Yin & McCowan 2004).

Morton, and later other researchers showed in a broad variety of terrestrial vertebrates that the structural changes of the vocal signals are generally depending on the inner state of the caller (Morton 1977). These "structural-motivational" rules state that more aggressive or dominant animals produce lower frequency and noisier signals, while subordinate or distressed animals' calls are high pitched and tonal (Figure 11).

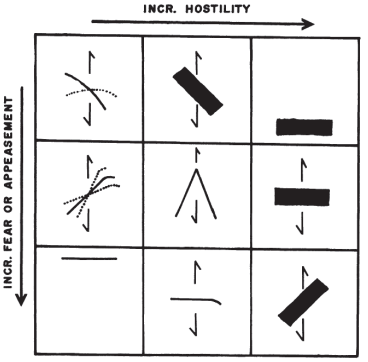

Figure 11. The visual representation of the structural-motivational rules. The frequency, tonality and contour of the pitch changes in aggression (horizontal) and fear (vertical) scales based on. (Morton 1977)

These structural regularities can be found in several bird species, Canids, deer and also in humans. It is assumed that the current state of the nervous system affect the sound production by influencing the muscle tenseness in the larynx or the vocal tract affecting the spectral structure of the calls (Titze 1994). Besides this frequency coding, in

the time domain the calling rate and duration can also carry emotional colouring of the signals (Pongrácz et al. 2006). Higher nervous activity can be in positive correlation with the rate and capacity of breathing, that determine the calling rate and intensity and provide information about the condition of the caller (McComb 1991; McElligott & Hayden 1999). The combination of the calling rate and duration can inform the receivers about urgency, especially in mobbing and alarm signals (Blumstein & Armitage 1997). Longer calls together with high rate of calling show the urgency of the situation, for instance they relate to the closeness of the spotted predator (Manser et al. 2002). High rate of calls evoke stronger sympathetic nervous reaction generating an inner state of 'fight or flight' in the receivers.

Vocalizations in several species like canines, primates, and humans too (Fox 1970; Drahota et al. 2008) co-occur with facial expressions (e.g. retraction or protrusion of the lips). These modifications of face shorten the vocal tract and result in vocalizations which signal a ‘shrunk’ apparent body size. This may be advantageous when indicating avoidance of fight. In contrast protrusion with closing the teeth have the opposite effect (Hauser et al. 1993; Harris et al. 2006; Chuenwattanapranithi et al. 2008).

Functional referential calls

In animal communication as we saw in the previous sections, most of the vocal signals convey information about the signaller itself, like physical, physiological parameters, emotional states (Evans & Marler 1995). However, since the first evidences were provided by Cheney and Seyfarth in the 80's, a growing number of species have been shown using calls with external reference (Seyfarth et al. 1980). Most researchers refer to referential communication if there is evidence that the emitted signal is independent from the inner state of the signaller. As we cannot have any direct information about the exact inner state of the animals, we are only able to assess or assign it from the context of signalling. Also the actual and the later behaviour of the signaller and the receivers of the signal can provide indirect information about the changes of the inner state (Seyfarth & Cheney 2003). In order to circumvent this problem animal signals which are seemingly referring to external things or events are called functionally referential signals (reviewed in: Macedonia & Evans 1993).

Alarm and food calls work typically as functionally referential signals because they provide information of presence of a predator or quality of food. Cheney and Seyfarth found that the vocalizations of vervet monkeys (*Chlorocebus pygerythrus*) refer to three

types of different predators (Seyfarth et al. 1980). Each type of predator has a distinctive predatory strategy. Hawks, eagles and other aerial predators attack from above, terrestrial predators such as leopards (*Panthera pardus*) approach their prey on the ground in the shade of the bushes, and snakes crawl close in the grass on the ground. Different predatory behaviour requires different flight strategies that exclude each other. In case of a ground predator monkeys are safe on the thinnest branches of the trees, while the densest bushes offer cover from birds of prey. In case of snakes it is not necessary to run up the trees because they move relatively slowly, but it is important to spot and keep an eye on them. Thus if a snake is close, the vervets stand straight and follow the movements of the predator in the grass (Struhsaker 1967). If a vervet makes the wrong move then it can be easily caught by the predator, and the same happens if it reacts too late. It is important to emphasise that during the development of individuals the range of evoking stimuli is narrowed by learning processes. Infants at earlier age emit also upon noting harmless animals of the same category. The sensitivity to the predatory species within a category emerges as they grow in parallel the relevant escape behaviour that has to be learned socially at earlier age (Seyfarth & Cheney 1980).

Thus these alarm calls are specific to the evoking stimulus; they are contextually independent, and independent also to a lesser degree from the inner state of the signallers, because any of the predators elicit a startle response in the monkeys. This latter is supported by observations showing that the closeness of the predator, the exposure, and the urgency does not affect the quality of the alarm call (Cheney & Seyfarth 1990). An alternative hypothesis is that the three alarm calls reflect different motivational state, that is, the vervet monkeys may have distinct types of fear responses. However this may be a less parsimonious explanation in this case. In recent years researchers have found other acoustic communication systems for which functional referentiality has been implicated (e.g. putty nosed monkey (*Cercopithecus nictitans martini*): Arnold & Zuberbühler, 2006; ravens (*Corvus corax*): Bugnyar, Kijne, & Kotrschal, 2001; chimpanzee (*Pan troglodytes*): Crockford & Boesch, 2003; fowl (*Gallus gallus*): C. S. Evans & L. Evans, 1999; red squirrels (*Tamiasciurus hudsonicus*): Greene & Meagher, 1998; suricata (*Suricata suricatta*) Manser et al., 2002; Diana monkey (*Cercopithecus diana diana*) Zuberbühler, Noe, & Seyfarth, 1997). Furthermore, in suricatas the predator specific alarm calls evoke similar behavioural response like in vervet monkeys but these calls convey urgency information also. This suggests that the graded and referential information in animal

signals are not exclusive (Manser 2001). Alarm calls in most species provide also cues about the signaller' size or identity (Cheney & Seyfarth 1988).

In conclusion functional referentiality depends on (1) structurally discrete and context specific calls which could be different call types or extremes of a graded continuum within one signal type (Fischer 1998; Rendall et al. 1999). Further it is important that (2) the eliciting stimuli belong to a coherent category (e.g. the signal of the aerial predators in vervet monkeys are evoked by dangerous hawks or eagles, but not by not dangerous vultures or crows). (3) Independently from the context the signal should evoke adequate behavioural response in the receivers in the presence (or absence) of the predator.

Not just alarm calls but food calls can fulfil these criteria, as they are typically used in the presence of edible objects as external referents. In chicken (*Gallus gallus domesticus*), roosters call a specific signal when they find a piece of food and the hens are also present, that respond with approach and looking for the signalled food. This reaction can be explained with two alternative models: (1) the food call signals the presence of the food for hens as a functional referential call, or (2) it provides social information such as low probability of aggression. In a playback experiment hens exclusively looked down to the ground searching for food when food calls were presented, while contact calls or alarm calls did not evoke such response suggesting the first model (Evans & Evans 1999). However in this case we cannot find different call types or coherent categories but the hens' reaction seems to confirm that the calls refer to external objects. More straightforward examples can be the food calls of primate species, where two different call types are associated with different quality foods (Hauser & Marler 1993). Besides the primates, other birds like ravens also seem to use food calls functioning referentially (Bugnyar et al. 2001; Slocombe & Zuberbühler 2005a).

There are three possible ways by which information can be coded in referential signals (Blumstein 1999a): (1) time code: one single call is produced sequentially (e.g. Morse code). This is typical for the dance language of the honeybees, but there is no evidence yet for such a mechanism in the vocal domain (Michelsen 1999). (2) The call intensity and other acoustical parameters may be varied on a continuum, and it can be categorized into two or more distinct subtypes that are context specific. Baboon barks grade from tonal to noisier, and noisier ones are used as alarm calls, while tonal ones are contact calls (Fischer et al. 2001b). The urgency is encoded in these calls by the call intensity; however this alone does not provide enough information about predators. (3)

Each evoking stimulus category has a distinct call type like in rhesus (Hauser 1998), lemurs (*Lemur catta*) (Macedonia 1990), prairie dogs (*Cynomys gunnisoni*) (Kiriazis & Slobodchikoff 2006) and chickens (Evans et al. 1993).

Question of honesty

Honesty has been a central question among researchers of animal communication. Do animals provide honest signals about themselves or they have the ability to manipulate the others in order to increase their own fitness (Dawkins & Krebs 1978; Zahavi et al. 1997)? As usual, there is no clear answer. In some circumstances only honest signalling is possible or the evolutionary race-of-arms resulted in a case of honest communication. The main problem in honest signalling is that how such system can be safe from cheating. According to Zahavi's handicap principle, the honest signals have to be costly to avoid the emergence of cheaters (Zahavi et al. 1997). Although more recent models showed that not the cost of signalling maintains a honest call system, but the potential cost of cheating (Számadó 2011). For example in a contest situation if the weaker individual signals itself stronger, in a real fight it will definitely loose. One could ask, in the light of this, what is happening in the case of the aforementioned red deer roars? The idea of proximity risk can solve this discrepancy (Számadó 2008). Proximity risk says that in fight context there is a distance threshold for honesty between the contestants. Within this distance the threat displays are reliable, while further away cheating can be present and evolutionary stabile. Red deer roars are typical long distant signals emitted from out of sight, therefore can be potential material for unreliable signalling. Also several other circumstances can maintain honest signalling, such as shared interest between the signaller and the receiver, when the signal is in strong relation with some physical parameters (indexes), individual recognition or social punishment of cheaters (reviewed in: Számadó 2011).

In other situations the species still can have the opportunity to bluff, lie and cheat. Two main types of manipulation can be differentiated: (1) withholding signals (not communicating in the presence of significant object or on occasion of a significant event), (2) misinforming (the individual communicates something that does not match with the reality) (Fitch & Hauser 2003). The second form seems to be the more abundant thus we concentrate on this, although there is a bias in the possibility of observation one or other, because it is far harder to find evidence for the first form (Cheney & Seyfarth 1990).

The act of misinforming has several preconditions. First, the signal has to be context specific, and second, individuals have to react to the signal in a stereotypic or consistent

way. Third, the signaller has to bear the ability to use the signal independently from its original context. The first two criteria concur with the criteria of functional referentiality, and the third flexibility makes misinforming possible. For the occurrence of functioning misinformation, it must happen rarely to avoid habituation of the receivers, it must have low cost on the deceived, and the testing of signal reliability must have a high cost. Alarm calls are perfect candidates for used as deceptive signals (Whiten & Byrne 1988). For example, the formosan squirrel (*Callosciurus erythraeus*) uses a post-copulatory call that is acoustically indistinguishable from the alarm call given to terrestrial predators (Tamura 1995). The neighbouring males react to the call as it was an alarm call and flee from the vicinity, thus this way the caller can reduce the possibility of secondary copulations and sperm competition, hereby benefit from the deception. In capuchin monkeys (*Cebus apella nigritus*) subordinates tended to emit a false terrestrial alarm call in the absence of predators to attempt usurpation of the food if dominant individuals monopolized a rich food resource (an elevated table with food provided by the experimenters) (Wheeler 2009). However, capuchins when hearing a terrestrial alarm call in such food related contexts, respond with a similar vigilance to spot the potential predator as in other contexts, but with less escape response. This seems as they are being suspicious about the possible deception (Wheeler 2010).

Misinforming about physical attributes is another possible way to deceit others. In contest situations showing a larger size may be a useful strategy, but from the evolutionary point of view it is especially important that testing the bluff should be costly for the receiver of the signal such as losing in the contest for the signaller (Goodenough et al. 2001b). A classic example for this bluffing behaviour is the case of the mantis shrimps (*Stomatopoda*). These stomatopods defend their burrows vigorously by showing their weaponry. They have two raptorial appendages that is able to strike on the opponent with a surprising strength and speed (it can break out the glass wall of an aquarium), and they present the colourful spots on the inner side of their appendices by spreading them to the side. This threat signal acts as an honest signal because the distance between the two appendages correlates with body size, and smaller individuals retreat when come across a larger one. However, when newly moulted, their cuticle is soft, and they lose the ability to strike effectively, and become extremely vulnerable. In spite of this softness, moulted individuals still use their threat signal in competitive contexts and threat away smaller, unmoulted individuals. Importantly, there is no clear sign of being moulted, so there is a high risk of injury when an individual tries to test the reliability of the others (Adams &

Caldwell 1990). In the vocal domain, the size exaggeration of the red deer can be considered as an attempt of similar misinforming behaviour (see above).

Vocal communication of the Canidae family

The Canidae represent a remarkable family (Figure 12) of the Carnivores covering almost the full spectrum of social systems from less-social, monogamous foxes (*Vulpes sp.*) to highly social, gregarious wolves and wild dogs (*Lycaon pictus*) (Fox 1975). Accordingly, their communicative repertoire, including vocalizations shows high variability along with the social spectrum (Fox 1971a; Cohen & Fox 1976). One species of this family, the subject of our dissertation is especially interesting for us because of the long shared evolutionary history with humans.

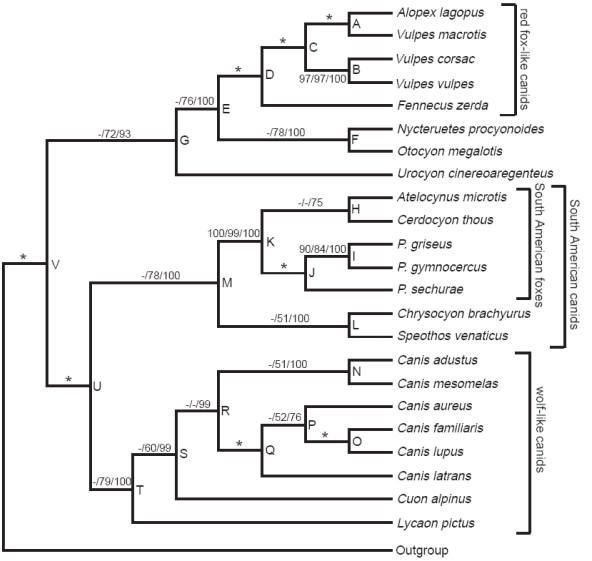

Figure 12. The consensus tree of the Canidae family based on six nuclear loci (from Bardeleben et al. 2005)

The vocal repertoire of the Canids

The earliest and still the most throughout comparative studies on canid vocalizations were made in the 1970's (Tembrock 1976; Cohen & Fox 1976; Lehner 1978). These studies aimed to record and characterize all vocalizations from different contexts and compare them among canid species. Subjects of these sound recording were mainly the close

relatives of the wolves and the family dog. Also these species are highly social and can be a good subject of collecting sound recordings from distinct behavioural contexts.

In these early works researchers recorded vocalizations, observed their spectrographic peculiarities, described and categorized them on the basis of various acoustical parameters. However, the nomenclature is somewhat different in these papers, we will attempt to summarize here all the described major types in the literature mainly based on the analysis of Cohen and Fox (1976) on several canid species and Schassburger's (1993) work on wolf vocalizations. We do not include to our list species specific calls (like rumble call of the wild dogs) and apply the grouping of Schassburger based on the call's tonality.

Tonal sounds

- ***Whine:*** short, cyclic high pitched vocalization, the fundamental frequency varies between 400 and 2000 Hz. There is an elongated variant ("*undulating whine*") that is characterized by modulated frequency. This can be considered as the fused cycles of whines. Another elongated subtype is referred as *whimper* with slight frequency rise and fall at the beginning and the end (Tembrock 1976) but with a relatively stable fundamental frequency. Wolves and wild dogs emit it during friendly greeting. *Cry* is similar to the whimpers, but it has a wider spectrum, more additional noise and stable fundamental frequency without the rise and fall at the beginning and the end (Tembrock 1976). Wild dogs use it as a begging signal, or for social solicitation (Robbins 2000). One interesting peculiarity in whines is the high occurrence of nonlinear phenomena (Volodina et al. 2006) that is typical in frustration vocalizations. Extremely high (6-8000 Hz) secondary fundamental frequency can appear in whines (referred as *whistle-whine* by Schassburger). These can be emitted parallel as biphonations, but as independent whines too. The latter case sometimes categorized as *squeals* like in dholes (*Cuon alpinus*) (Volodin et al. 2001). Whines are rare in foxes and are emitted only when in pain, while dogs, wolves and coyotes (*Canis latrans*) use them in different contexts, especially along with stressed, submissive and contact seeking behaviours.
- ***Yelp****:* short (0.2-0.3 s), rapidly following vocalisation with a fundamental frequency between 400 and 800 Hz and with clear harmonic structure with a

bark like Christmas tree shape. It can appear in playful and painful situations but the play yelps are more bark like, while the stress yelps are more whine like (Bleicher 1963). According to Tembrock, (1976) adult foxes use bark-like yelps as defensive calls, however these calls are more like barks in the classification of Cohen & Fox, (1976). Wild dogs and dholes call an elongated version of yelps (called *squeals*) in greeting contexts (Robbins 2000; Volodin et al. 2001). A short, high frequency tonal call similar to yelp is also described in dhole (called *yap*) (Volodin et al. 2001). Bleicher suggested that the yelps develop from whines and evolve into barks, thus these vocalizations are acoustically in between these two calls (Bleicher 1963) and appears only in pups, except in dogs where adults still use this call. A shortened, tonal variation of yelp is called as *yip* in coyotes and jackals (*Canis aureus*). Lehner did not differentiate it from the yelp and reported that coyotes use it in high intensity submission and alarm or startle (Lehner 1978).

- ***Screams:*** elongated (3-4 s) sound, can be cyclic with a high fundamental frequency (1200-2000 Hz), and contains more noise. It can be heard in pain and submission in coyotes, wolves and dogs, while foxes use it more elaborately in several different contexts.

Intermediate sounds

- ***Moan:*** Intermediate long calls (0.1-1s), lower fundamental frequency (80-600 Hz) and fast frequency modulations differentiate these from whines. According to Schassburger, moan is intermediate between growls and whines, thus several subtypes can be differentiated along the tonality axis from *whine-moans*, *harmonic moans* to *noisy-* and *growl-moans*. They appear in several different contexts, probably in ambivalent inner state, and function as frustration and contact seeking calls or the noisier variants as playful and greeting signals. Also mentioned in wild dogs (Robbins 2000).
- ***Barks****:* Very short calls (less than 0.5 s) with a variable fundamental frequency (150-500 Hz) and noisiness. The fundamental frequency shows a curved shape, and the energy concentrated in the lower parts of the spectrum causing a typical Christmas tree shape (Feddersen-Petersen 2000). Foxes use individual barks as threat, while wolves and dogs barks cyclically or in fused bouts in alarm and threat, but dogs use it in several other different contexts.

Foxes also emit short bouts as a long distance call (*staccato, wow-wow* and *yodel barks* – Newton-Fisher et al. (1993)) Coyotes and jackals also use it when alarmed, in territory defence and just jackals in hunting coordination (Lehner 1978).

- ***Howl****:* Very long or extended vocalizations with sometimes undulating fundamental frequency varying between 150 and 2000 Hz (Cohen & Fox 1976). Frequency alternations are stronger in the beginning and end, with jumps occasionally in the middle (Tembrock 1976). These calls are specific to wolves, coyotes and dogs, used as group vocalization, contact seeking, defence or greeting contexts, perhaps these can be acoustically distinguishable as it is in coyotes (Lehner 1978). The g*rumble* is a moderately noisy howl with typical pulsing similar in the *staccato* call and also used in conflict situation (Volodin et al. 2001) by only the dhole, while the *hoo* is a subtype with clear tonal fundamental frequency around 1000 Hz (Robbins 2000), it is the wild dogs' long distance contact call similar as lone howls in other canids.
- ***Groan****:* Spectrally and acoustically similar to moans. It has a relatively low or medium fundamental frequency (250-450 Hz) modulated periodically, with modest noisiness. these calls emitted by dog pups and adults in acute distress, pain or sickness (Bleicher 1963).
- ***Cough****:* Short, growl-bark like sound emitted with closed mouth. Used by all canids possibly as a warning signal. Foxes usually emit it in bouts (Newton-Fisher et al. 1993).
- ***Grunt****:* Low frequency (85-200 Hz) short, harmonic or slightly noisy calls emitted during relief, comfort or pleasure (Bleicher 1963). This vocalization is not known in foxes.

Noisy sounds

- ***Growl****:* Short and repetitive or elongated, broadband, low frequency (80-300 Hz) vocalizations with low frequency variability and high amount of noise but with still visible harmonic structure (Riede & Fitch 1999). Foxes use it as warning or threat and defence, while dogs, coyotes and wolves additionally emit in greeting, group vocalization and play (Yeon 2007). One subtype emitted with closed mouth through the nose is called *snort*.

- ***Snarl****:* Low frequency (fundamental at 140-170 Hz), broadband, long calls (1-2 s) with strong noisiness (Tembrock 1976). The most extreme variant of growl emitted with opened mouth, retracted lips and clearly visible, exposed canines.
- ***Woof****:* a low amplitude, soft single bark with low fundamental frequency (90-120 Hz) used as mild threat or alarm used by wolves, coyotes and foxes (Lehner 1978) and also mentioned in dogs by Feddersen-Petersen, (2000), but it is possibly just a subtype of barks used as a modest threat or warning.

Atonal sounds

- *Pant:* Specific, noisy, forced in- and exhale in dogs and foxes, during play soliciting and greeting (Cohen & Fox 1976).
- *Tooth snap:* This sound is used in wolves, coyotes and dogs, in play, and agonistic contexts. The sound is generated by the fast closing and collision of the teeth.

As we saw, the canine communicative repertoire is very rich, moreover canids also use intermixes of these clear classes that can multiply the possible calls that can be used in a sophisticated communicative system (Cohen & Fox 1976).

The dog bark

The domesticated dog and the grey wolf are so close genetically to each other that some taxonomists differentiate them only on subspecies level. However vocalisation wise, the only domesticated member of the Canidae family differs markedly from its closest relative. Dogs emit barks at higher frequency and use them in a wider range of contexts (Cohen & Fox 1976). Some researchers have assumed that barks in dogs are only excitement signals, and the visual signals are the really important ones during communication (Bleicher 1963). Others argued that barks were hypertrophied during domestication and become structurally more complex and variable. This, together with their contextual diversity might suggest that this hypertrophy is due to relaxation of the selection pressure for silence (Cohen & Fox 1976), thus lacks information content (Coppinger & Feinstein 1991). This idea is supported by the observations that feral dogs barks similarly rare and in limited contexts as wolves (Boitani & Ciucci 1995). More recently the possibility was raised that barks function as a mobbing signal, and the frequent disturbance in human environment evokes mobbing behaviour in dogs and cause the elaborate barking behaviour (Lord et al. 2009)

However, Tembrock has already suggested that barks may contain contextual information: he found that play and aggression bark sequences differed in their temporal structure (Tembrock 1976).

More recent studies showed that dog barks are not only diverse, but acoustic subtypes of barks can be reliably linked to social contexts (Feddersen-Petersen 2000; Yin & McCowan 2004). Moreover, Pongrácz and colleagues found that human listeners are able to categorize dog barks correctly by the context (Pongrácz et al. 2005), and that they can judge the assumed emotional state of the dog which is associated with the particular situation e.g. whether the dogs were aggressive, fearful or playful (Pongrácz et al. 2006). Humans seemed to use in this assessment the general rules of Morton, because their judgement was strongly affected by the frequency and temporal structure of the bark sequences. Interestingly, humans were not successful when they had to differentiate barks of individuals (Molnár et al. 2006), nonetheless barks seems to vary not just contextually and along the signaller's inner state but show individual differences (Molnár et al. 2008). As humans can reliably recognize the context of dog barks, the possibility of context dependent information of these signals emerged also. However, functional referentiality has been usually investigated in the case of communication among conspecifics, and until now, only two studies examined the role of barks in dog-dog communication. These studies show only that dogs are able to differentiate barks emitted in distinct social contexts (Maros et al. 2008) as well as between individuals (Molnár et al. 2009b), but not the use of the possibly encoded information.

Dog growls: a more conservative kind of vocalization

While diverse types of barking seems to be predominantly a dog specific vocalization, comparative vocal analysis in canids, including dogs, so far found little interspecific differences in the function of growls (Tembrock 1976; Cohen & Fox 1976; Yeon 2007). In contrast to barks that became hipertrophysied during domestication and may have been modified by the humans' selective work, and may play a role in dog-human communication (Pongrácz et al. 2010), growls appear to be more conserved. Wolves and dogs growl in three different situations: (1) offensive threatening in social conflict, (2), defensive threatening such as guarding food, or (3) during social play (for review see: Yeon 2007). The fact that growls are emitted in two markedly different affective states (agonistic and playful) suggests that they carry the opportunity to act as graded, functionally referential, or maybe both kinds of signals.

Moreover, dog growls seem to be excellent candidates for studying size communication. Due to their broadband and noisy frequency structure, their formants are relatively easy to locate in the spectrum. Taylor et al. (2008) has already shown that defensive growls of the dogs convey information about the size of the signaller because both fundamental frequency and formant dispersion is correlates negatively with the body size (see also: Riede & Fitch, 1999) and humans were able to use these cues to assess the size of the growling dog (Taylor et al. 2008). Also dogs that show extreme size variance between and within breeds use growls in competitive contexts, where size communication can be crucial. However there is no data yet on how other dogs are able to use this indexical cue, and whether their ability to extract information from the growls and how their acoustical structure can carry this information about the caller or its context.

Aims, questions

Our studies aimed to explore the communicative aspects of dog growls. We wanted to test whether growls convey information about the context, the caller's inner state and physical attributes. As most dogs of the Western world predominantly live within close social relationship with humans, we wanted to test the communicative content of growls both in case of dogs and humans as receivers.

Experiment 1 and 2

In the last decade several studies show that some of the graded vocalizations are also context specific (Fischer 1998; Rendall et al. 1999). According to this view, these call types have the potential to carry context specific information if the graded calls can be sorted along an axis of one or more acoustical parameters into several distinct subtypes with relatively small overlap, and these subtypes can be wedded to different behavioural contexts (Marler 1976). Moreover, the two or more differing calls should basically represent the same inner state. For example, a study on captive chimpanzees (*Pan troglodytes),* showed that two subtypes of a food call, the "rough grunts", can be distinguished based on the quality of the discovered food (Slocombe & Zuberbühler 2005a). One can hypothesize that a chimpanzee gets basically into similar inner states when he/she finds either kind of food. Thus, when the animal produces different calls in the case of the two types of food, the quality of the vocalisations might represent the quality of the food item, and not the different inner states of the chimpanzee.

Several studies already show that besides alarm and food calls, signals used in various social contexts may be also functionally referential (Rendall et al. 1999; Crockford & Boesch 2003). Slocombe & Zuberbühler, (2005b) reported that chimpanzees use different subtypes of screams during agonistic encounters, one of which can be linked to the role of aggressor and the other to the role of the victim. They suggest that these calls can function referentially because out of sight third party individuals seem to use these signals to decide whether to interfere or not.

Until now, the vocal communication of large terrestrial carnivores has not been investigated from the aspect of functional referentiality, even though it is known that species of the *Canidae* family use different vocalizations in a wide range of social contexts (Tembrock 1976; Cohen & Fox 1976; Lehner 1978).

According to Hauser's three criteria for functional referentiality, we first conducted an acoustical analysis of growls recorded from three different situations (two agonistic and one non-agonistic situation): 1) the subject dog guards his food from an unfamiliar conspecific (food guarding growl - FG), 2) an unfamiliar human approaches threateningly to the subject dog (threatening growl - TS) and 3) the subject dog plays with the owner a tug of war game (play growl - PL). We hypothesised that growls recorded from different social contexts will differ in their acoustical parameters. Based on the motivation-structural rules of Morton, and the findings of Pongrácz and Molnár in the acoustical differences between dog barks, we can assume that the playful and the agonistic growls will differ in their pitch, tonality and rhythmicity, while the threatening stranger and food guarding growls will be less different due to their similar inner state and context. We also expect difference between the two agonistic growls due to their different contextual background.

Second, we conducted a playback experiment in a semi-natural situation. The playback experiment corresponded to the food guarding situation. The aim of the experiment was to investigate whether growls recorded in the food guarding situation have the potential to evoke the expected behaviour (retreat) from the receiver in the physical absence of the sender.

Third, in the same experiment we expected to see differences in the receiver's behaviour when hearing agonistic growls or the play growl in the food guarding situation. We hypothesized that the food guarding growl, which was recorded in a similar situation (Food Guarding), has the most deterrent effect on receiver dogs, while the other two types (TS and PL) have no or only a weaker effect. The difference between the effects of the situation specific growl (FG) and the two other types of growls (TS, PL) would indicate

that the dogs' behaviour is influenced dominantly by the context specific information. Alternatively, if no difference is found between the effects of the three growl types, or if the two agonistic growls (FG and TS) had an equally strong deterrent effect on the receiver dogs, then the receiver's behaviour was probably mostly influenced by the affective content of the vocalisation.

Experiment 3

Additionally we were curious about how dogs are able to use the formant coded size information within growls. Several theoretical and field studies have shown that animals estimate the physical characteristics of the opponent before starting costly fights (Enquist & Leimar 1983; Arnott & Elwood 2009). Often such estimations are based on visual displays before starting a fight (e.g. red deer: (Clutton-Brock et al. 1979), elephant seals: (Cox 1981) and cichlid fish: (Enquist et al. 1990)). However, ritualized displays of body size are often considered dishonest signals since the signallers may appear larger than they are (e.g. parallel walking of deer, fur erecting, spreading fins, erecting gills, etc. (Maynard-Smith & Harper 2003)).

In human speech, the shape and change of the formant frequencies are responsible for the differentiation of speech sounds (Lieberman & Blumstein 1988), but also carry indexical cues about the callers' sex and body size (Smith & Patterson 2005). Early anatomical studies suggest that the manipulation of the vocal tracts' shape and size is uniquely human. However, Fitch and Reby (2001) have shown that besides humans, several mammalian species, including pigs, goats, monkeys and dogs, are able to actively change the characteristics of their vocal tract during vocalization. When a species gains such a trait, it gives an opportunity for dishonest signalling, which can lead to an evolutionary arms race between signallers and receivers modifying their ability to precisely perceive the exact size of the caller (Fitch & Hauser 2003). In summation, perception of formants and formant spacing can be an important cue for judging size of conspecifics, as several studies have shown in humans (Feinberg et al. 2005), the rhesus monkey (Fitch 1997) and dogs (Riede & Fitch 1999; Taylor et al. 2008).

In the light of these, our food guarding growl recordings can be obvious candidates for studying this phenomenon. Thus to answer the question of size communication in growls, we conducted a second playback study with an intermodal presentation of growl with additional visual information to test the auditory size assessment abilities of dogs. We

assumed that if dogs are able to accurately assess the size of the growling dog they will match this information with the visual modality.

Experiment 4

Not just intra-, but interspecific communication can be studied with playback methods (Shriner 1998). For example, in studying dogs' vocal communication, their close relationship with humans is a considerable advantage, that gives an opportunity of getting not just behavioural responses of conspecifics, but the humans' opinion about the callers and their vocalizations (Pongrácz et al. 2006; Taylor et al. 2008). Humans are extremely vocal creatures, and they use an elaborated vocal repertoire during communication. Human language carries both referential information and cues about the speaker's inner state (Sauter et al. 2010). This latter information is encoded in the physical structure of the emitted vocalisations by the changes of tempo and energy of the fundamental frequency ("prosody") (Hawk et al. 2009), and may be also in the formant structure (Chuenwattanapranithi et al. 2008). Not surprisingly humans are able to rely on these acoustic features to assess others' inner state during social interactions. In humans specific brain regions are involved in this process which are different from those that are responsible for speech perception (Schirmer & Kotz 2006; Fecteau et al. 2007). A neuroimaging study by Belin et al. (2008) found that in humans the same brain centres are responsible for the processing of animal (cat and rhesus macaque) and human vocalisations with a negative valence. This may indicate the presence of ancient homologue brain structures in humans for processing non-verbal acoustic content.

Thus, as growls are emitted in contexts where the possible inner state of the dog is presumably have opposing valence (playfulness-aggressiveness) and possibly also carry contextual information, thus for last we played back the growls for humans, following the methods of the bark studies of Pongrácz and Molnár. We aimed to explore how humans are able to recognize the context of the growls and the assumed inner state of the dogs to explore further the possible contextual and affective information content of growls. We hypothesised here that similarly as with barks, humans are able to recognize contexts of growls, and can attribute matching inner states to them.

Experiments

Experiment 1: Acoustic analysis of dog growls (Faragó et al. 2010b)

In the first step, it is necessary to collect the studied vocalizations to explore their potential to convey acoustical information. Thus, we have to record these sounds in the natural environment of the animal and record the context of these spontaneous vocalizations, this way we can get a picture of the vocal repertoire of the studied animal, and can have assumptions if some of these vocalizations can be connected to given contexts, inner states or physical parameters of the subjects. If focusing at some specialized vocalizations like alarm calls, or signals used in social contexts, a better option to evoke these calls by putting the subject animals into an artificially generated and controlled context, in which it is known that the animal will vocalize. This way we still cannot have information about the inner state of the subject during vocalization, but we can control and standardize the conditions of the recording, which highly ease the comparison of the different types or sub-types of a vocal signal. This approach was used for example to elicit alarm calls in marmot species (*Marmota sp*.) (Blumstein 1999b), and likewise in the studies of the dogs' barking behaviour of the last decade (Yin & McCowan 2004; Pongrácz et al. 2005), or dog growls (Riede & Fitch 1999; Taylor et al. 2008), and was used in our studies too.

Fitch & Reby (2001) reported that some mammals are able to modify their vocal tract to change the acoustic pattern of their vocalizations. Fitch's cineradiographic observations also showed that dogs' vocal tract is highly modifiable, which can serve as anatomical basis for the diverse vocalization potential of dogs (Fitch 2000). Also this ability can provide enough potential acoustical variability within a call type such as the growl to use it in so different social contexts. In order to test this, we conducted acoustical analysis on our collected growl samples, and compared several basic parameters like fundamental frequency, formant dispersion, tonality and growl length that will potentially vary along the three contexts.

Materials and Methods

Sound recording

For collecting growl samples from dogs, we applied a similar method for sound recording like Csaba Molnár in his bark studies with a difference in that we made our recordings in

closed environment to minimalize external noise, and used different social contexts to evoke growls from the subjects.

In our studies, growls of 25 adult dogs (for details see Appendix, Table 5) were recorded between 2007 and 2010 in three different social contexts using a Sony Digital Audio Tape Walkman (type: TCD-D100) with a directional microphone (type: ECM-MS907) and later H4next Handy digital recorder. All growls were recorded in a silent, 5.5 m x 4 m empty room. The contexts were the following (presented in the order of recording, see also Figure 13, Photo A, B and C).

Threatening stranger (TS) Photo C: The dog was held on a leash by the owner who stayed behind the dog. The owner was not allowed to touch or talk to the dog. The experimenter was crouching near to the dog to keep the microphone as close as possible (15-30 cm). A stranger (25 year old male) started to approach the dog, slowly, silently and keeping tense eye contact with the dog. If the dog did not pay attention to the stranger, he tried to draw the dog's attention towards him with coughs or foot taps (see details in Vas et al. (2005)). This slow approach was finished when approximately 10 growls were recorded, or just before the stranger reached the dog. After the recording, the stranger started to talk to the dog in a friendly manner and play with it.

Tug of war game (PL) Photo B: The owner started to play with the dog using a rug, a ball or another dog toy. The experimenter stood near the dog to keep the microphone as close as possible (15-30 cm). The play was continued until approximately 10 growls were recorded.

Food guarding (FG) Photo A: The dog was kept on a leash, and he/she was given a large, meaty bone. We waited until the dog started to chew on the bone. At this point, an unfamiliar dog was let to approach the chewing dog (both dogs were on a leash and controlled by their owners). The experimenter held the microphone as close as possible (15-30 cm). The unknown dog stayed and tried to reach the bone until approximately 10 growls were recorded. We took care that the situation did not escalate into a fight, and no dogs were harmed.

Figure 13. Pictures taken from the three sound recording contexts A, Food guarding, B, Playing Tug-of-war, C, Threatening stranger (Photos taken by András Péter)

For editing the sound recordings, we used Adobe Audition 1.5 software (Adobe Systems Incorporated). The entire recording of each dog was copied on the computer with 16-bit resolution and 44.1 kHz sampling rate for editing. The recorded sequence of growls was split into separate growls (average length: FG: 1.53±0.2s; TS: 1.36±0.18s; PL: 0.68±0.09s**)** from which we selected those with the least background noise. Then we made copies of each growl and normalized them (to -6dB) for the acoustical analysis. After this, from the intact growls, we constructed different growl sequences (16.5±7.8 second long) from four or five growls originating from the same situation and from the same dog. These sequences were saved as different files, and labelled with a number code.

Sounds

We have used the growl recordings of 20 dogs collected between 2007 and 2009 for acoustical analysis (see Appendix: Table 7).

Acoustical analysis

We used an advanced linguistic software called Praat (Boersma & Weenink 2001) to perform the acoustical analysis. This software is used more and more widely in bioacoustics because of its flexibility and custom-required programmability. During the acoustical measurements Praat macro scripts (see supplementary file) were used to extract automatically the following parameters:

- growl length (duration) (L),
- fundamental frequency (F0: cross-correlation method, 125ms time window, 20-2000Hz frequency range),
- formant frequencies (F1-5: burg method, 25ms time window, maximum frequency 6000Hz, maximum five formants),
- formant dispersion (dF: calculated as given by Riede & Fitch (1999)),
- standard deviation of formant dispersion (SDF: calculated as given by Riede & Fitch (1999)).
- harmonic to noise ratio (HNR: cross correlation method, 10ms time step, minimum pitch 75Hz).

The fundamental and formant frequencies were visually checked on a random sample of 30 growls, and the automatic assessment was found accurate. We calculated these parameters for each growl sample of the 20 subjects (average number of samples: 11.24±5.85), and for further analysis we used the mean of the parameters of the growls originated from the same dog and from the same situation.

Statistics

All the statistical analyses were conducted in SPSS 15 (SPSS Incorporated). The acoustic parameters of the three groups were compared with linear Mixed Effects Models (Fixed factor: type of growl, Random factor (repetition): subject) and Tukey HSD post hoc test ($P < 0.05$).

Results

The comparison of the mean growls from 20 dogs showed marked difference in the acoustical parameters of growl duration, fundamental frequency and the formant dispersion (Mixed Effects Models: L: $F_{2,18} = 4.103$, $P = 0.04$; F0: $F_{2,18} = 4.355$, $P = 0.034$; dF: $F_{2,18} = 9.638$, $P = 0.002$). Post hoc tests showed that FG and TS growls did not differ from each other (for detailed parameters see Appendix: Table 6). However, we found that play growls are at least half time shorter (Figure 14) as well as higher pitched than the other two, agonistic growl types (Figure 15). Also the formant dispersion of the play growls was lower than in the other two growl types. This shows that, the average frequency step between the formants were smaller in the play growls, that is the play growls' formant frequencies were relatively lower (Figure 16).

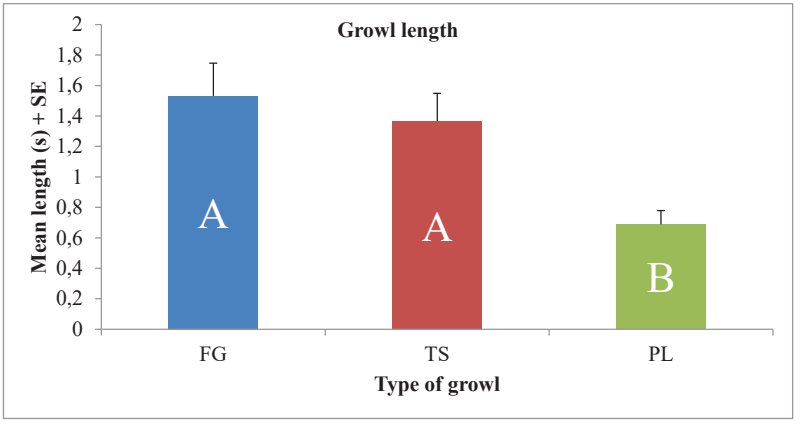

Figure 14. The mean length of the individual growls in the three contexts. The different letters indicate significant difference.

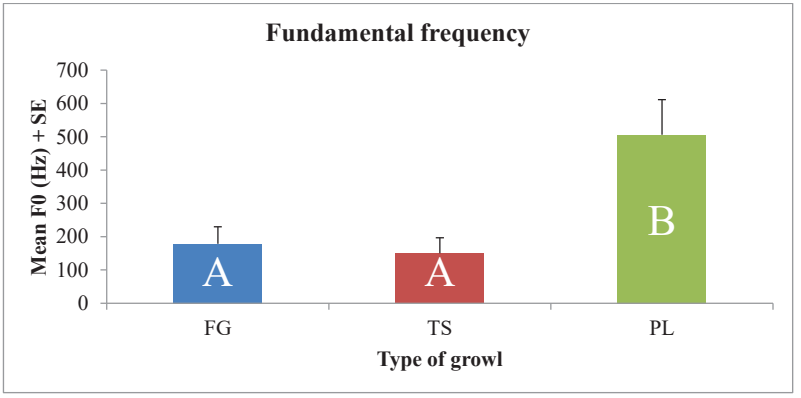

Figure 15. The mean fundamental frequency of the individual growls in the three contexts. The different letters indicate significant difference.

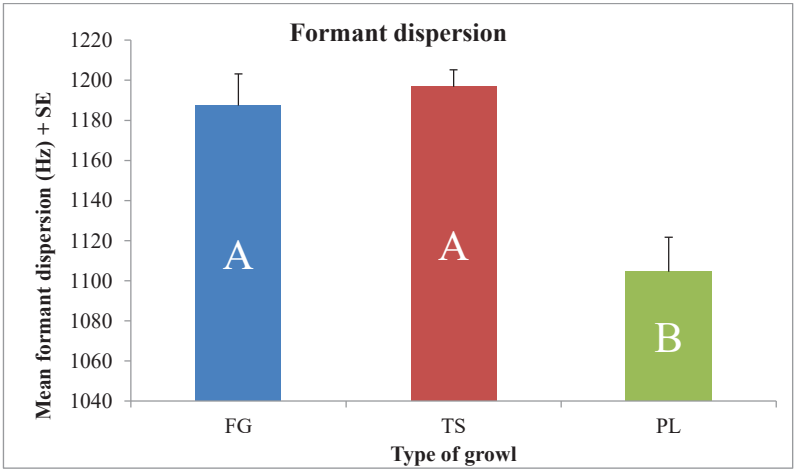

Figure 16. The mean formant dispersion of the individual growls in the three contexts. The different letters indicate significant difference.

The average height at withers and weight of the dogs whose growls were used for analysis did not differ significantly between the three groups (Mixed Effects Model: height $F_{2,14} = 0.906$, $P = 0.426$, weight $F_{2,14} = 0.906$, $P = 0.426$).

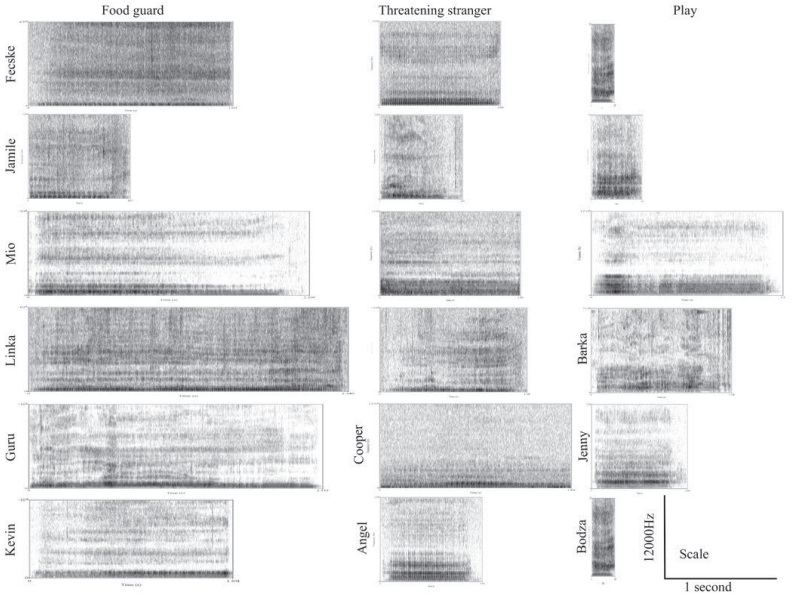

Figure 17. Spectrogram samples of growls from the three different contexts.

Discussion

The two most important findings of our acoustical analysis were the differences between the fundamental frequencies and the formant dispersions of the play growls and the other two agonistic growls. We found that play growls have a higher fundamental frequency and lower formant dispersion than the growls originated from the two agonistic contexts. According to Taylor et al. (2008), growls with lower fundamental frequencies belong usually to larger dogs, so the deeper pitched growls in the agonistic situations (FG and TS growls) could indicate that the individual dogs appear 'bigger' with regard to body size than the same dogs giving a play growl. However, according to Riede & Fitch (1999) the formant dispersion indicates just the opposite, namely that when a dog is producing a play growl, he/she shows itself bigger than during producing an agonistic growl. We hypothesize that dogs modify their growls, and thus use growling either to communicate aggressive intention in an agonistic situation, or the lack of this during play. Additionally, the tension and position of the vocal folds can be modified affecting the fundamental frequency, while pulling back the lips with open mouth can cause 'shortened' vocal tract and widely spaced formants, and lowering the larynx or vocalising with lips closed on the

side of the mouth and open at the frontal part can cause the lowering the formants and shorten the formant dispersion (Fitch 2000). This can explain why play growls have lower formant dispersion, because dogs during tug-of-war game hold the tug in their mouth and growled with closed teeth, which also lengthens the vocal tract. However, we would need more detailed physiological observations like cineradiographical studies to reveal the mechanism behind this formant difference.

Interestingly, we did not find any significant difference between the HNR of the three growl types. HNR refers to the noisiness of the sounds (Riede et al. 2005) and it was one of the main variables which discriminated aggressive and non-aggressive barks for human listeners (Pongrácz et al. 2006). This phenomenon might be attributed to the fact that barking represents an unusually diverse cluster of vocalizations emitted in various situations (Feddersen-Petersen 2000), while growls are used in narrower range of contexts.

We have to admit that our sample of growls in present form, due to the low subject number is not applicable for a detailed acoustical analysis (for example permuted Discriminant Function Analysis – see Mundry & Sommer (2007)) which could show that the different type of growls can be differentiated by their acoustical parameters or not. However the results of the following experiments will show that growls from different agonistic contexts have distinguishable effect on dogs' behaviour, indicating indirectly that there are subtle context-specific acoustic differences among these growls.

Experiment 2: Bone is mine – affective and referential aspects of growls (Faragó et al. 2010b)

By the means of recording and analysing vocalizations we can have an impression about the production side of the vocal communication. However, without studying the perceptual part of a communication system we cannot see the whole picture and will not be able to understand it. During playback studies, we can test the behavioural responses of the recipients of the vocal signal; we can test what information can be transferred by the vocalization in relatively controlled circumstances.

In the first study, we found that dogs' growl in markedly different contexts, and these growls show acoustical differences too, but whether the growls can be interpreted as functionally referential signals however, is not clear. Our study is the first which aims to test the contextual specificity of these vocalisations of dogs. Instead of going out to the field (the home of the owners) and trying to test the behaviour of dogs during playbacks, we chose to set up our studies in the lab, where we can control the environment and the speaker-amplifier and camera system is available. While habituation-dishabituation studies with dog barks observed only the attention and orientation behaviour of the dogs, and got information only about the discriminative ability of the dogs (Maros et al. 2008; Molnár et al. 2009b), in our study about growls we put the subjects into a semi-natural context to see their spontaneous reaction.

Materials and methods

Subjects

Subjects were 41 adult pet dogs recruited from the databases of the Clever Dog Lab (http://www.nc.univie.ac.at/index.php?id=14571) in Vienna and the Family Dog Project in Budapest (http://kutyaetologia.elte.hu/). We used subjects from different breeds and ages (minimum 1 year old. For detailed information see Appendix: Table 7. Five subjects had to be excluded from the experiment, because they did not show interest in the bone.

Experimental setup

The playbacks were conducted in a 3.6 m x 14 m (Vienna), and a 3 m x 6 m (Budapest), silent room (Figure 18). The setup was basically the same in the two rooms. We placed a dog cage covered with a blanket in the room and hid a speaker system in the cage. In Vienna we used a 3.1 speaker system (type: Medion sound system), whereas in Budapest a

Genius type speaker pair, with Technics amplifiers (type SE-A909S and SU-C909U) was used. We placed a fresh, cooked calf bone (20-25 cm long, approximately. 0.5 kg heavy) on the floor, at 20 cm distance from the cage. We used one bone only once, according to hygienic considerations and to avoid the effect of the smell of an earlier subject's saliva.

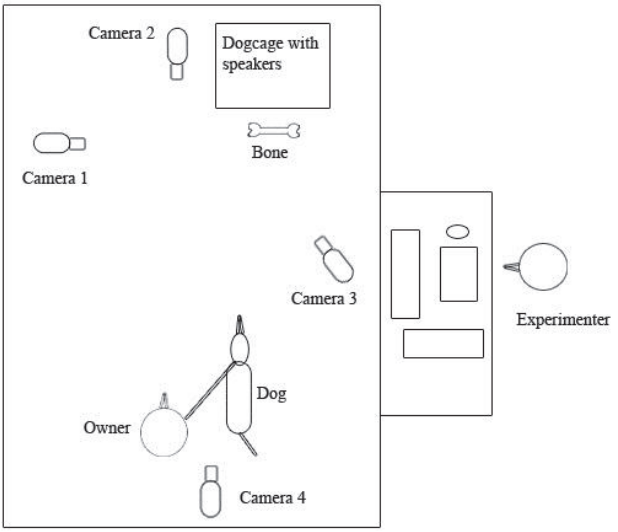

Figure 18. Schematic drawing of the experimental set-up.

During the experiment, the owner was alone with the dog in the room, while the experimenter with an assistant who operated the sound playbacks, stayed in the adjacent room following and recording the events in the experimental room through a webcam connected to a computer outside the room. In Vienna, we used a camera (type: JVC HDD camcorder GZ MG20E) to record the subjects behaviour, which was fixed on a tripod, and placed at the posterior third of the room. For the live observation we used a webcam (type: Creative Live! Cam, Vista IM) fixed on the anterior wall of the room, at 1.9 m height. In Budapest, we used a quad camera system both for recording and for the live observation. The live picture was recorded on the computer with Virtual Dub software. The same computer was also used to playback the sound sequence with Adobe Audition 1.5 software.

Procedure

Each dog participated in only one experimental trial, because in a pilot study we found a very strong order effect when we conducted within-subject design, even if there were several days between two tests with the same dog. In an experimental trial, the subjects heard a growl recorded from one of the three social contexts: TS, PL or FG. Dogs were randomly assigned to one of the three experimental groups and 12 dogs were tested in each group. Growl sequences for the playback experiments were chosen randomly from the pool of sound recordings from the 20 dogs. In sum, 11 dogs' growls were used for playbacks (see Table 1 and Table 2).

Growls from the same dog in the same context were played back to no more than two subjects and even these growls consisted of different sequences. Thus, we used six dogs' growls in each group. Therefore, no two subjects heard exactly the same sequence. Moreover, we measured the sound pressure level of each sequence prior to the experiment, and corrected the volume manually, so that the playbacks' volume did not differ between the experimental groups (TS: 72.9±1.1dB, N = 12; PL: 73.9±1.7dB, N = 12; FG: 72.9±0.6dB, N = 12; Kruskal-Wallis test: $\chi^2_2 = 1.414$, P = 0.493).

At the beginning of a trial, the owner entered the experimental room with the dog on leash. First, the owner was instructed to lead the dog around the room to let the dog get acquainted with the room and to discover the presence of the bone. The dog was allowed to sniff at the bone only from a distance of about 20 cm, but was not allowed to touch it. After the dog sniffed at the bone, the owner was told to lead his/her dog 3 m away from the dog cage and stay there. After a knocking signal from the experimenter, the owner released the dog from the leash. The owner was instructed to refrain from talking, touching or looking at the dog. The experimenter and the assistant observed the dog on the monitor, which was connected to the camera. If the dog did not approach the bone within 2 minutes, the trial was stopped and the dog was excluded from the analysis.

If the subject approached the bone within 5 centimetres (in essence, just before it could reach it) the assistant started to play back the assigned growl sequence. If the dog left the bone (withdrew its nose more than 5 centimetres from the bone) playing the sequence back was stopped after the actual growl was over. If the dog did not approach the bone again within the next one and a half minutes, the trial was terminated. If the dog approached the bone again, the experimenter continued to play back the growl sequence. If the dog did not leave the bone, the whole sequence was played back a maximum of three times, and if the dog was still chewing the bone, the trial was terminated. If the dog took

the bone and moved away from the cage more than 20 centimetres, the trial was also terminated. If the dog left, but went back to the bone repeatedly, the trial was continued and the growls were played back each time the dog approached the bone. The trial was finished if the dog left the bone unattended for more than 1.5 minute, or started to chew on the bone in front of the cage, or took the bone away from the cage.

Behaviour analyses

The behaviour analyses and coding were conducted with Solomon Coder beta 080510 software (developed by András Péter, at ELTE TTK Department of Ethology, Budapest). Each trial was divided into two parts for the analysis: First, the pretest phase, which started with the closing of the door after the owner had entered the room with the dog and ended when the dog was released from the leash. With the release of the dog the test phase started, which ended when the experimenter entered the room to stop the experiment.

The frequency (contacting the bone, leaving the bone) and the duration (licking, chewing, taking away the bone) of the dogs' bone oriented behaviour were recorded and analysed. The descriptions of the behaviour elements are in Table 1.

Name of behaviour Element	Definition
Contact with the bone	The dog touches the bone with his/her nose or mouth
Licking the bone	The dog repeatedly touches the bone with her/his tongue
Chewing the bone	The dog takes the bone into his/her mouth and chews it without taking it away more than 20 centimetres from the original position
Taking away the bone	The dog stands near the bone, lowers the head and grabs the bone with mouth and moves it more than 20 centimetres away from the original position
Leaving the bone	The dog steps away at least with the two front legs from the bone and her/his head moves more than 50 centimetres away from it

Table 1. The definitions of the coded behaviour elements.

During the analysis we measured the following variables: (1) Latency of the second contact with the bone (the maximum latency was 90 seconds (see procedures)); (2) Latency of leaving the bone for the last time (latencies were measured from the first contact with the bone, the maximum was defined as 180 s according to the maximum duration of the trials - time elapsed from the first contact to the end of the trial); (3)

Frequency of contacts with the bone (number of contacts divided with the total time elapsed from the first contact to the end of the test trial); (4) Time percentage of handling the bone (total duration of licking and chewing the bone divided with the total time from the first contact to the end of the test trial).

Statistical analyses

All the statistical analyses were conducted in SPSS 15 (SPSS Incorporated). Before the main analysis we examined the possible effect of the different testing locations (Table 2). Since none of the behaviour elements were significantly different, we pooled the subjects from Vienna and Budapest for further analysis. None of the behavioural data met the criteria homogeneity of variance and some groups did not show normal distribution, thus we used Kruskal-Wallis test with Bonferroni post-hoc test ($P<0.05$). Tests were all two tailed.

	t_{34}	*P*
Latency of second contact with the bone	-0,45	0,656
Latency of leaving last time the bone	0,763	0,451
Contact frequency	0,86	0,396
Time percentage of handling the bone	0,791	0,435

Table 2. The comparison of the behaviour elements of subjects from the two testing site.

Results

Individual responses

After hearing the growl for the first time, 11 of 12 dogs in the FG group withdrew from the bone within 15 s, whereas only two of the 12 dogs from the TS group and four of the 12 dogs in the PL group withdrew within 15 s. Seven dogs from the FG group did not approach the bone again within the next 90 s, while only one in the TS and one in the PL group stayed permanently away from the bone. Figure 19. shows the distribution of individual responses of the dogs in the three groups. That marked difference that most of the dogs in the FG group left the bone, and only one ate it, but in the TS group most of the dogs took the bone away, or ate it suggests that FG growls had stronger deterrent effect on the behaviour of dogs than TS growls. While in the PL group more dogs left the bone than in the TS group shows that these growls evoked ambivalent responses from the dogs.

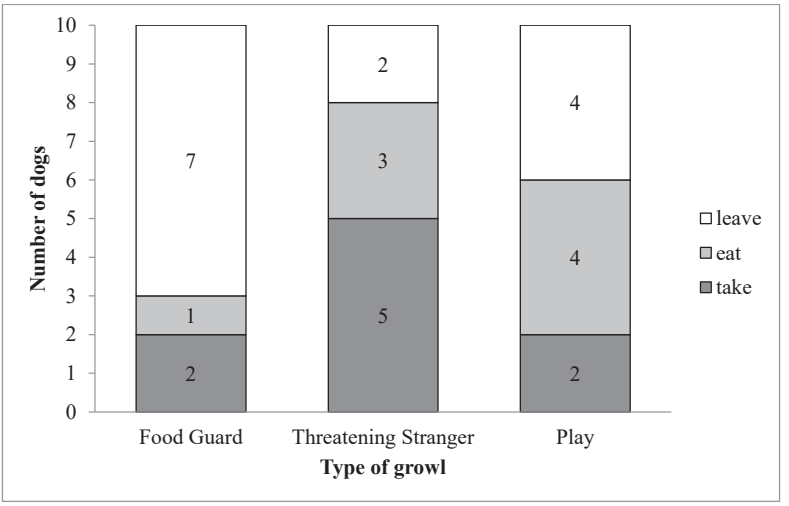

Figure 19. The distribution of individual reactions among the three groups. The numbers in the columns show the number of dogs showing the given behaviour.

Latencies of contacting with the bone and withdrawing from the bone

We found significant differences among the groups in the latencies of reapproaching the bone after being scared away the first time by the playback (Kruskal-Wallis test: $X^2 2 = 11.34$, $P = 0.003$). The Bonferroni post hoc test revealed that dogs in the FG group reapproached the bone significantly later than the dogs in the other two groups (FG vs. TS: $P = 0.008$; FG vs. PL: $P = 0.006$). There was no significant difference between the TS and PL groups ($P = 1.0$) (Table 3, Figure 20.).

The latencies of leaving the bone for the last time also differed in the three groups (Kruskal-Wallis test: $X^2 2 = 10.098$, $P = 0.006$). The post hoc tests showed that latencies of leaving the bone for the last time were the shortest in the FG group (FG vs. TS: $P = 0.002$; FG vs. PL: $P = 0.032$), and there was no difference between the latencies of the TS and PL groups ($P = 0.887$) (Table 3, Figure 21.).

Table 3. Descriptive data of the analysed behavioural variables.

	Food Guard				Threatening Stranger				Play			
	Mean	**Median**	**Interquartiles**	**Standard deviation**	**Mean**	**Median**	**Interquartiles**	**Standard deviation**	**Mean**	**Median**	**Interquartiles**	**Standard deviation**
Latency of second contact with the bone	57	90	23.18/90	35,86	161,2	6.38	2.75/15.81	32,99	131,8	4,12	2.06/10.68	33,19
Latency of leaving last time the bone	65,08	8	0.75/155.18	77,89	20,16	180	180/180	51,63	18,68	180	52.12/180	70,88
Contact frequency	0,32	0.01	0.01/0.05	0,05	0,06	0.05	0.04/0.09	0,03	0,04	0.04	0.01/0.06	0,03
Time percentage of handling the bone	13,55	0.78	0/32.77	23,64	33,96	26.94	8.55/63.51	30,29	28	23.35	2.26/48.41	26,83

Durations and frequencies

There were no significant differences among the groups either in the frequency of contacting the bone (Kruskal-Wallis test: $X^2 2 = 4.704$, $P = 0.095$) or in the total percentage of dogs handling the bone (Kruskal-Wallis test: $X^2 2 = 4.33$; $P = 0.115$) (Table 3).

Alternative explanations

We compared the relative playback durations (length of growl playbacks divided with the numbers of contacts in a given trial). We did not find significant difference between the groups (Kruskal-Wallis test: $X^2 2 = 5.818$, $P = 0.06$). It means that the differences were not caused by the different playback lengths in the three experimental groups.

We also tested whether the differences between the body sizes of the growling dogs caused the behavioural differences between the experimental groups. Thus we made three groups from the subjects by the weight of the growling dogs (small size: 13-20kg, N=5, mean: 17.28±0.56; middle size: 29-30kg, N=3, mean: 29.6±0.16; large size: 32-35kg, N=3, mean: 34.16±0.32), and we compared the behaviour of these groups with Kruskal-Wallis test. There were no significant differences (latency of the second contact: $X^2 2 = 5.394$, $P = 0.07$; latency of leaving the bone last time: $X^2 2 = 2.764$, $P = 0.251$; frequency of contacting the bone: $X^2 2 = 1.994$, $P = 0.369$; percentage of handling the bone: $X^2 2 = 1.584$, $P = 0.453$), hence we can reject the alternative hypothesis that the bigger dogs' growls had stronger effect than the smaller ones.

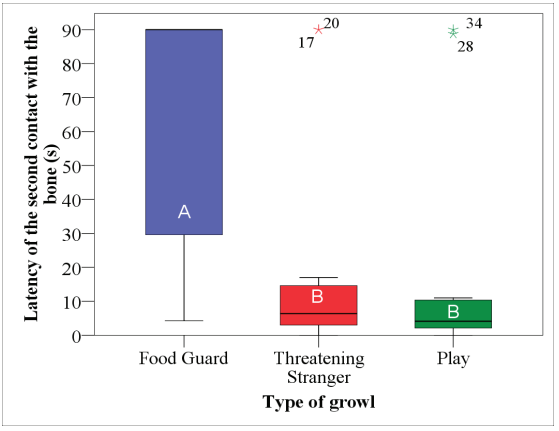

Figure 20. The medians of the latency of reaproaching the bone in the three groups. The columns indicate the interquartiles, the whiskers the lowest and highest non-outlier values, while the stars represent the outliers. The different letters in the boxplots indicate significant difference.

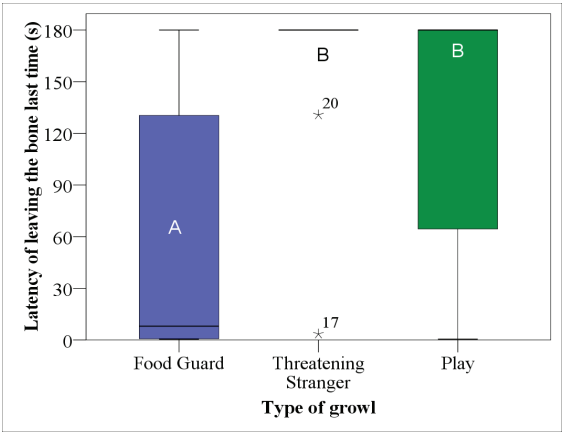

Figure 21. The medians of the latency of leaving the bone in the three groups. The columns indicate the interquartiles, the whiskers the lowest and highest non-outlier values, while the stars represent the outliers. The different letters in the boxplots indicate significant difference.

Discussion

Dogs avoided the bone mostly when they heard FG growls, while the PL and TS growls were less effective deterrents. If the dogs would have reacted differently only in the case of the PL growls then the different inner states of the signaller (play vs. defensive aggression) could explain the differences in the receivers' behaviour. However, the results of the playback experiment suggest that the dogs differentiated markedly the FG and TS growls. If we assume that the growling dogs were in similar inner states in the two agonistic (FG and TS) situations, then the effect of FG growls could be interpreted in the functionally referential framework. These results are also in agreement with Hauser's arguments (1998, see above), because the contextually most relevant vocalisation had the largest effect on the behaviour of dogs. PL and TS growls remained ineffective as deterrents, because the subjects probably did not process them as contextually appropriate.

An alternative hypothesis regarding functionally referential communication is the possible existence of different inner states behind the different signals (Owren & Rendall 1997). In our case it would mean that dogs have markedly different inner states in the FG and TS contexts, and these are reflected directly in the acoustic structure of the different vocalizations. Our acoustic analysis contradicts somewhat this hypothesis, because we found only subtle or no differences between the TS and FG growls when analysing the wide range of the well-established acoustic parameters (Riede & Fitch 1999; Pongrácz et al. 2006; Taylor et al. 2008). One would expect more distinct signals if the context would

be less important than the signals' acoustic structure. However, we cannot exclude the possibility that dogs differentiate TS and FG growls by attending to other more subtle acoustical features, which we did not measure in this study. One possible difference can be caused by that the dogs growled on the threatening human always with closed mouth, but during defending their food they often showed their teeth and pulled back their lips.

In summary, there were strong differences between the reactions to the two agonistic growl types, when only one of them was appropriate in the given context. Based on the approach of Hauser (Hauser 1998) and Cheney and Seyfarth (Cheney & Seyfarth 2007), it would be possible to assume that dog growls are functionally referential. However, we should also take into consideration simpler explanations. It is obvious that dogs can discriminate between the FG and TS growls. Studies, like the one of Fischer (1998) showed that slight acoustic differences between particular calls are enough for Barbary macaques (*Macaca sylvanus*) to associate them with different meanings. It is highly possible that many of the subject dogs learned it earlier, how and with what kind of consequences does a food guarding, or a threatened dog growl, hence they can show appropriate behaviour, independently from the intentionality of the signaller. Unfortunately, subtle differences between inner states of dogs during various agonistic situations are difficult to measure, as well as communicative intentions, so whether growls are referential or not will remain somewhat a philosophical question. In the future bigger emphasis should be laid on more detailed acoustical analysis of the differences between similar, but relevant signals, such as FG and TS growls.

Experiment 3: Dogs' expectation about signallers' body size by virtue of their growls (Faragó et al. 2010a)

As we saw in the first two studies, growls seems important for communication because they carry context specific information (Faragó et al. 2010b). Moreover, Taylor and co-workers have shown that growls recorded in agonistic contexts carry indexical cues because formant spacing correlated with body size (Taylor et al. 2008). In addition, humans were sensitive to this indexical cue (Taylor et al. 2008) and dogs showed a different pattern of explorative behaviour in response to different formant-manipulated growls (Taylor et al. 2010a). However, it is still unknown whether dogs are able to match size related acoustical information with visual size information.

In order to investigate this question, we used the preferential-looking paradigm. This method is broadly used in child psychology and psycholinguistics presenting static pictures or videos accompanied with sound playbacks to pre-linguistic infants to explore their ability to match cross-modal information (e.g. Walker, 1982). This intermodal perception "is the perception of unitary objects or events that make information simultaneously available to more than one sense (Bahrick & Hollich 2008)". The whole paradigm is based on this ability. The rationale of this is that individuals prefer to look at visual stimuli that correspond to broadcasted auditory stimuli, instead of non-matching visual stimuli, when the subject is able to link together the two modalities. For example when we show a dog and a cat picture simultaneously to the subject, and play back a dog bark or a cat meow, and the subject recognize the pictures and the sound it will look first and longer at the matching picture because it will have an inner representation of the source of the sound and expect the sound coming from the direction of the adequate picture (Guihou & Vauclair 2008). This method can be well adapted to nonhuman animals too to test several questions like individual recognition (Sliwa et al. 2011), numerical abilities (Jordan et al. 2005), size perception (Ghazanfar et al. 2007) and communication (Ghazanfar & Logothetis 2003). Majority of such studies were made on primates but other species were can be good subjects (Seyfarth & Cheney 2009). For example horses individual recognition was successfully tested with cross-modal presentation (Proops et al. 2009), while Adachi et al demonstrated that dogs are able to match the sound and face of their owners (Adachi et al. 2007).

Therefore, in order to test if dogs could also match visual with acoustical information, we presented dogs with a choice of two pictures of the same dog, together

with a playback of a growl recorded in a food competition context. One of the two pictures was manipulated to show smaller or larger body size compared to the growling dog, while the other picture was matched in size to the growling dog. We expected, according to the other studies using the preferential-looking-time paradigm, that dogs would look longer at a picture of a dog that was matched to the signallers' body size and moreover, that no preference for either picture would be shown in control trials where either no size information was given in the acoustical domain (presentation of noise) or non-dog pictures were presented (shapes or cat pictures).

Materials and Methods

Subjects

Participants in our study were well socialized family dogs, recruited from the database of the Clever Dog Lab, Vienna. In total, 116 dogs participated in this study. Of these, 20 dogs had to be excluded from the sample due to technical problems (n= 11) and behavioural difficulties (n= 9), leaving a total of 96 dogs (male: 42, female: 56, mean age: 4.7±2.6 years) for analyses.

Experimental design

We used four different types of stimuli pairs during the experiment in a parallel intermodal presentation (see (Guihou & Vauclair 2008)):

Dog-Growl (DG) – matching modalities: presentation of dog growls with projection of dog pictures. (N= 24)

Dog-Noise (DN) – non-informative sound: presentation of Brownian noise with projections of dog pictures. This control was conducted to test if there was any effect caused by the size difference of the dog pictures on the looking behaviour of the subjects. (N= 24)

Shape-Growl (SG) – non-informative picture: presentation of dog growls with projection of triangles. This control group was used to test if the looking preferences of the dogs were nonspecific in that they would occur also with non-natural, unknown pictures. (N= 24)

Cat-growl (CG) – non-matching modalities: presentation of dog growls with projection of cat pictures. This control was conducted to test the effect of the nature of the pictures. (N= 24)

These latter two controls were specifically designed to test if possible positive results of the test condition was dependent on the nature of the visual stimuli or not. We tested 24 subjects in each of the four experimental groups. All dogs tested in the first three groups were familiar with other dogs. Moreover, subjects in the last group were also familiar with cats (once lived, currently living together or in regular contact with cats).

We used a between-subject design to avoid order effects and possible habituation of the subjects to the stimuli. Thus, each dog was tested only once with one sound sample paired with a visual stimuli.

Stimuli design

The auditory stimuli were chosen from a pool of pre-recorded food guarding growls from 12 different dogs and previously generated Brownian noise samples (Adobe audition 1.5). We used a food guarding growl because these growls are used against other conspecifics in an agonistic encounter. The physical parameters (height at withers, weight) of the twelve growling dogs were known (see Appendix Table 8). We used the growls of two groups of markedly different sized dogs (six "small dogs": shorter than 52 cm and six "large dogs": taller than 60 cm at the withers) for the experiments (Figure 22.). The growls in the two size categories differed significantly in their acoustic structure (parameters were measured with PRAAT). The formant dispersion, as mentioned in our introduction, is in strong negative correlation with body size, thereby acting as a reliable cue. The same was true in our growl samples: small dog growls had significantly higher formant dispersion than the large growls (mean dF - small: 2658±232Hz; large: 671±253Hz; unpaired t-test: **t=14.17; p<0.0001**). The fundamental frequency in most of the examined species shows only weak or no correlation with body size because the anatomy of the vocal folds is not affected by the skeletal and muscular conformation of the body (Fitch & Hauser 2003). However, different dog breeds show such high morphological diversity, that their vocal fold size and fundamental frequency can also correlate with the physical parameters of the body (Taylor et al. 2008). The two subsets of growls used in this study also differed in their fundamental frequencies: small dogs' growls had significantly higher fundamental frequency than the growls of the large dogs (mean F0 - small: 130±25 Hz; large: 90±253Hz; unpaired t-test: **t=2.83; p=0.018**) suggesting that the two acoustical parameters together can act as an indexical cue for the dogs in our study.

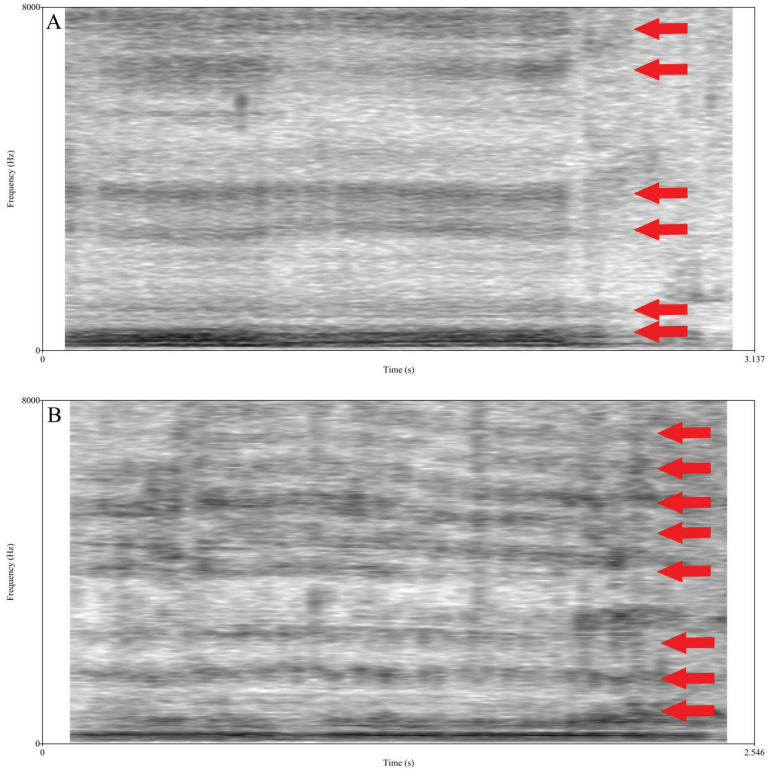

Figure 22. Two samples of a growl recorded from a small (A) and a large (B) dog. The red arrows indicate the formant frequencies giving a hint about the lower formant dispersion in the large dog.

The visual stimuli were digital photos of twelve different short-nosed dog breeds (DG, DN), twelve different cat breeds (CG) and twelve differently coloured triangle shapes with various internal angles (SG) (for some examples see Figure 23.). The pictures were presented in front of a homogenous black background. In each experimental trial, we showed the same two pictures, but one was adjusted to be life sized (the size of the projection was similar to the size of the actual growling dog - matching), while the other picture was 30% smaller or larger than the matching picture (resized). The projections' height was measured on the canvas from the ground to the withers in the case of dogs (22-83 cm) and cats (22-67 cm). The length of the vertical side was used for the height of the triangles.

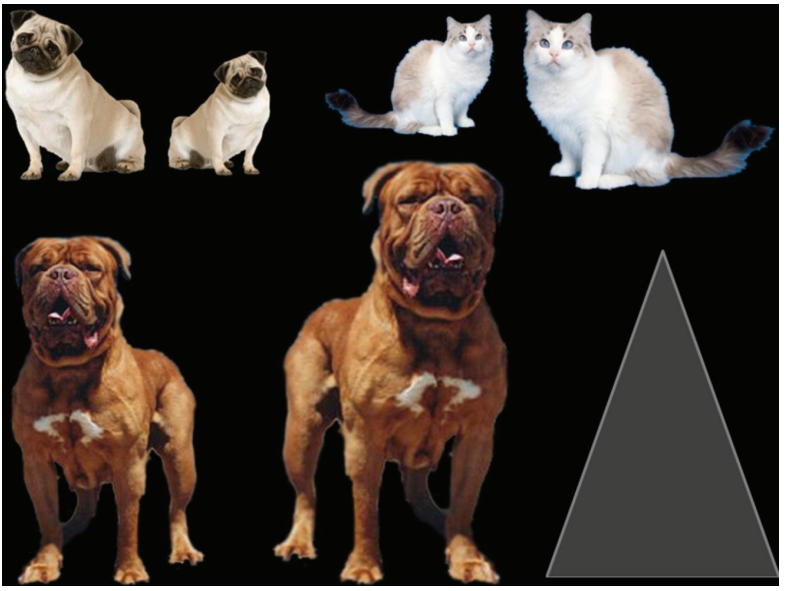

Figure 23. Examples of the used visual stimuli.

Within each group half of the subjects heard a "small dog" growl, and the other half heard a "large dog" growl. The side where the manipulated picture appeared was balanced within each group. Each growl was presented with the same pictures to two different dogs, but the size was differently adjusted: one dog saw a smaller picture paired with the matching sized photo, while the other one saw a larger picture paired with the matching sized photo. Thus, each growl was used two times for playbacks within each group, except in the Cat-Growl group. In the CG group we used only the six small dogs' growls, as cat pictures adjusted to match their size to the large dog growls would be unnaturally big. Thus, in the Cat group, each growl sample was used in sum four times, twice-twice with two different cat pictures (for the detailed set see Table 8).

Experimental set-up

A chair for the dog's owner was placed at one side of the darkened experimental room (5 m x 6 m) facing the canvas on which the pictures will be presented to the dog (Figure 24.). During the entire experiment, the owner was listening to music through headphones to prevent him/her from hearing the playbacks and influencing the behaviour of the dog unintentionally.

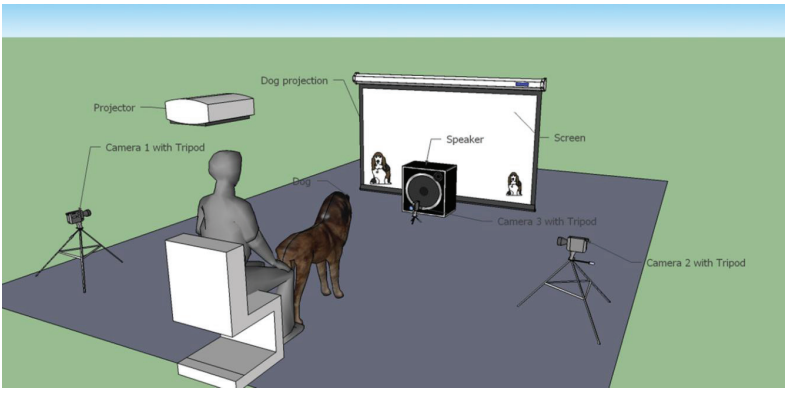

Figure 24. Schematic drawing of the experimental set-up.

The dog sat or laid between the owners legs, also facing towards the canvas (size: 2.3 m x 1 m) on which the two pictures will be shown (height of 30-95 cm depending on the size of the growling dog, 2.2 m apart). The pictures were presented by a projector which was positioned behind the owner at a height of 1.5 m. The speaker was placed on the ground in front of the canvas in the middle of the two presented pictures. We used four cameras to record the behavioural response of the dog. The first camera was put next to the projector (Cam1) recording the projection on the canvas. The second camera (Cam 2) – a zero lux camera that could record in low light density – was put in front of the speaker to record the dogs' response. Near the owner's chair, an infrared lamp was directed to the dog's face in order to lighten them up for the zero lux camera, without interfering with the projections. Finally, two wide-angle cameras with high light sensitivity recorded the events in the entire room (Cam3 and 4). We used Cam1's microphone for sound recording. For effective projection of pictures, the lights were switched off and the windows were covered with a curtain. The experimenter controlled the events from the neighbouring room via a closed-circuit video system and a PC was used for the picture presentation, audio play back and video recording.

The stimuli were displayed during the experiments as PowerPoint slideshows. Three slides were used: the first and the last were homogenous black, while the middle slide contained the pictures and the sound. The change between the first and second slide (stimuli presentation) was controlled by the experimenter, and was dependent on the behaviour and attention of the subject (see below for the exact criteria) while the sound sample was automatically played as soon as the visual stimuli appeared. The change

between the second and third slides (disappearance of the pictures) was automatic after 20 seconds. The volume of the different growls was measured from the location where the dogs sat and adjusted to the same level (65dB) prior to the experimental trials.

Procedure

Before the experiment, the owner was informed about the procedure and told his/her tasks during the experiment, although no detailed information such as the type of given stimuli was explained. Next, the dog and its owner entered the experimental room and the dog was allowed to explore the unfamiliar room for approximately 2 minutes. During the experiment, the owner was asked to sit on the chair with headphones on and listen to music on an mp3 player. The volume of the music was adjusted to a level that prevented the owner from hearing the sound playbacks.

The dog was sitting or lying in front of the owner on the floor, facing the screen (beginning body posture). The dog was not on leash and the owner was allowed to position it with their hands just before the appearance of the pictures. The experimenter switched off the lights and asked the owner to adjust the dog gently into the beginning body posture. Then he zoomed Cam2 at the dogs' face and started the video recording before finally leaving the room. During these events the first black slide was projected at the canvas.

After the experimenter left and the door was closed, the dog was in the beginning body posture for at least 10 seconds and the sagittal axis of the dog's head pointed to the centre of the canvas (looking to the middle), the experimenter then switched to the second slide which activated the growl. The projected pictures then appeared on the canvas at the two lateral parts of the lower side of the canvas for 20 seconds. During the projection of pictures, the owners were not allowed to look at the pictures and they were not allowed to talk or touch their dog. If the owner interacted with the dog during the projections, the data were not used for the analysis. After the projection of the pictures ended, the experimenter entered the room and the experiment was finished. If the dog did not orient towards the canvas during the growl playback, or did not look at either of the pictures, the trial was considered to have failed and the data of the subject were not used for the analyses.

Behaviour analysis

We measured the looking preference of the dogs. Their orienting behaviour and the time of looking at the two pictures during the projections were blind coded by an assistant (Nándor Takács) with a Solomon Coder (behaviour coding software developed by András Péter,

Dept. of Ethology, Budapest). Inter observer reliability was tested on a random subsample (10 dogs from each group) of recordings (average Cohen Kappa: 0.87). Orientation towards one side was defined as the angle between the sagittal axis of the dogs' head and the centre of one of the pictures, which is less than 10 degrees (looking at the picture), and the dogs' eyes fixed on that projected picture (Figure 25). We coded the behaviour from the beginning of the projection until the pictures disappeared to measure the latencies and the looking time, the maximum of these time variables was 20 seconds. We transformed the latency data to a binary variable to analyse which picture the subject glanced at first (1 – looking first at matching picture, 0 – looking at resized picture). We calculated the overall looking time (the sum of time looking at any of the two pictures) and the time ratio of looking at the matching picture (time of looking at one given picture divided with the overall looking time). The ratio of looking at the bigger picture and at the left picture was also calculated to test the possible effect of the size and the position of the picture independently from the acoustic stimuli within groups.

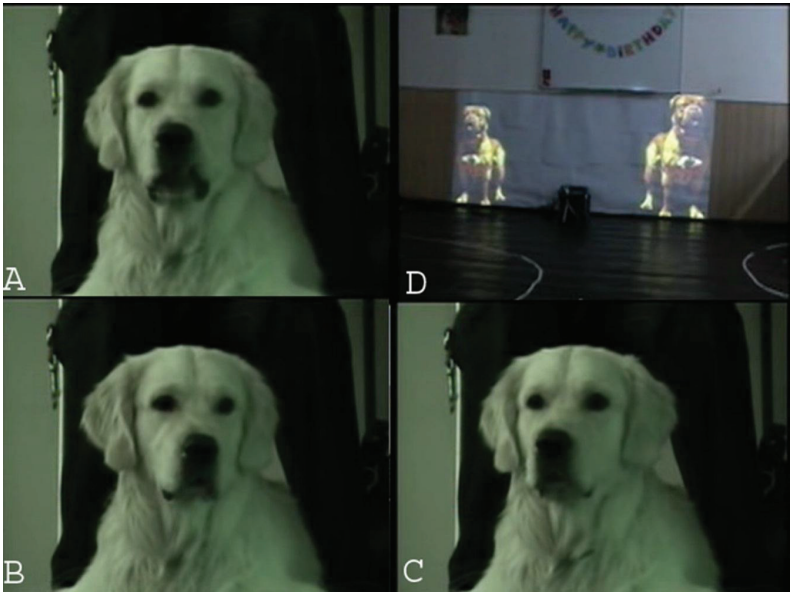

Figure 25. Examples of the looking reactions of a dog. A, looking at the middle at the start of the playback. B, looking at the left picture. C, looking at the right picture. D, the projected dog pictures with the speaker in the middle. The left (smaller) picture is matching with the actual growl playback. (Frames taken from the original experimental video footage.)

Statistical analysis

None of our behavioural variables were normally distributed, thus we used non-parametric statistical tests. To test within-group effects, we used exact Binomial and Wilcoxon matched-pair tests, for between group comparisons we used Kruskal-Wallis tests. To measure looking and side preferences, the looking time of the groups was compared to a hypothetical 0.5 median (no preference) with Wilcoxon signed-rank tests. All tests were performed with SPSS 15, except the looking preference tests and the Dunn post hoc tests, for these GraphPad Instat statistical software was used. For correcting on multiple measures, we applied FDR correction.

Results

The latency of looking at the matching and the resized picture showed that dogs in the Dog-Growl group looked earlier at the matching picture than at the non-matching picture. 20 of 24 subjects in this group looked first at the matching picture (two tailed Binomial test with 50% expected value: p=0.0015), while in the other three groups we found no significant difference from the expected 50% (Binomial test: DN (50%): p=1; SG (58.3%): p=0.541; CG (41.7%): p=0.541) (Figure 26). Moreover, the dogs in the test group also showed a strong looking preference towards the matching picture (Wilcoxon test: **p=0.0022**) compared to the control groups (Wilcoxon test: DN: p=0.9441; SG: p= 0.8334; CG: p=0.8774) (Figure 27).

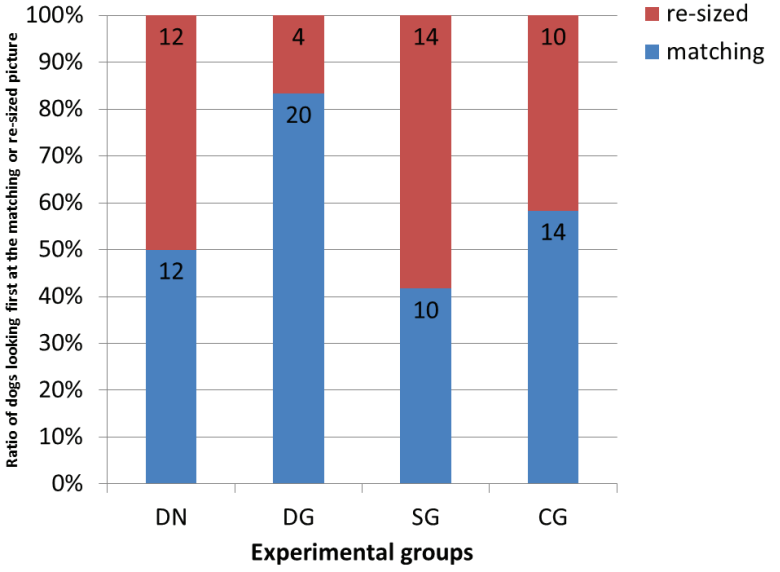

Figure 26. The distribution of the individual reactions in the four groups. The numbers in the columns represent the actual numbers of dogs looking first at the matching or the re-sized picture.

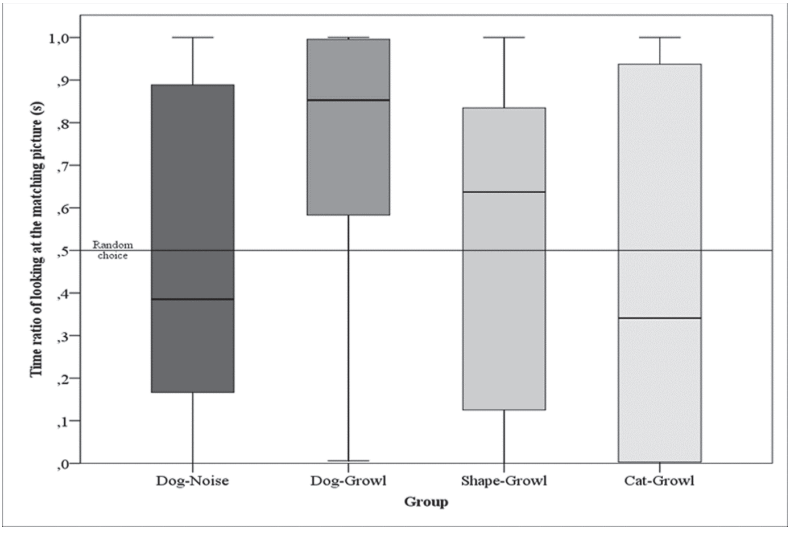

Figure 27. The looking preference of the dogs in the four experimental groups. The line at 0.5 time ratio indicates the lack of preference, the area above the line shows looking preference at the matching picture. The boxes indicate the interquartiles, the whiskers the lowest and highest non-outlier values.

When comparing the behaviour of the dogs in all of the experimental groups, we found that dogs spent more time looking at the stimuli when animal pictures were presented as compared to triangles (Kruskal-Wallis test: $\chi 2=27.003$; $p<0.001$; Dunn post hoc test: DN vs. DG: $p>0.05$; **DN vs. S: p<0.001**; DN vs. C: $p>0.05$;**DG vs. SG: p<0.01**; DG vs. CG: $p>0.05$; **SG vs. CG: p<0.001**), suggesting that the visual information conveyed by the pictures on natural objects was more interesting to the dogs than the artificial shapes (Figure 28).

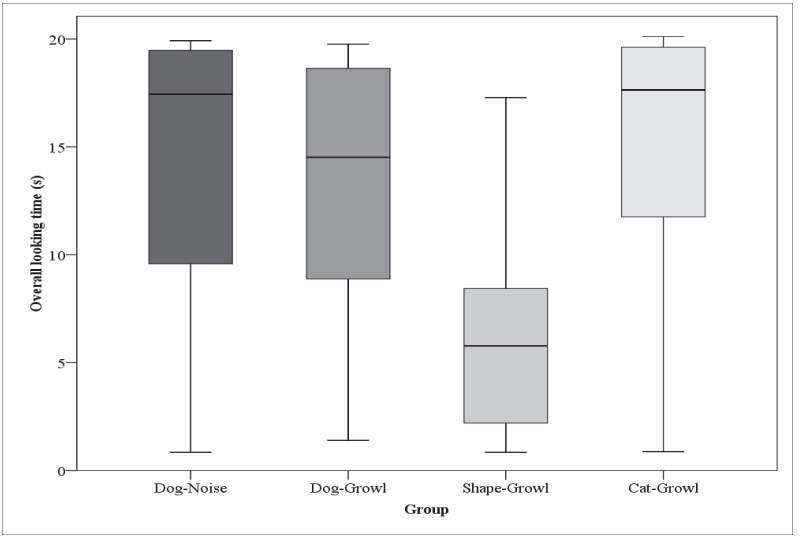

Figure 28. The medians of the overall looking time of the dogs in the four experimental groups. The boxes indicate the interquartiles, the whiskers the lowest and highest non-outlier values.

None of the groups showed significant preference towards the bigger picture (Wilcoxon test: DN: $p=0.0738$; DG: $p=0.8334$; SG: $p=0.4389$; CG: $p=0.4732$). This suggests that the visual proximity did not affect the subjects' looking behaviour (Figure 29).

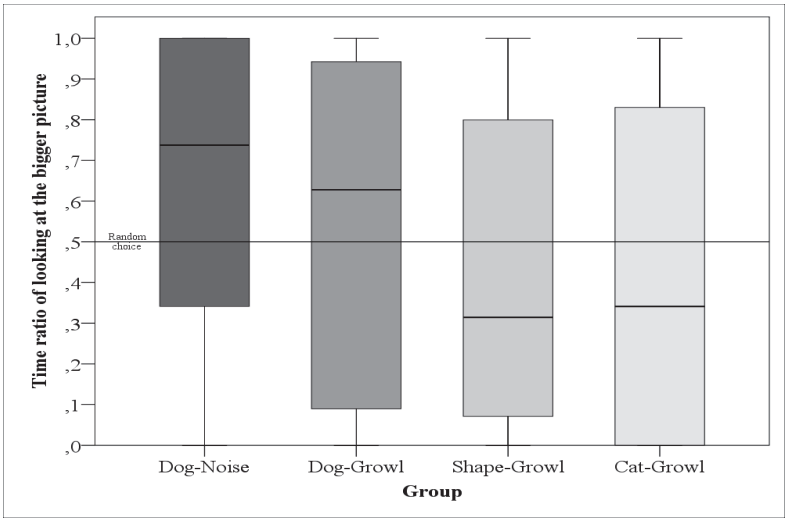

Figure 29. The dogs' looking preference towards the larger picture in the four experimental groups. The line at 0.5 time ratio indicates the lack of preference, the area above the line shows looking preference at the larger picture. The boxes indicate the interquartiles, the whiskers the lowest and highest non-outlier values.

Moreover we tested if other factors had an effect on the looking behaviour of the dogs, but neither the size of the growling dog (Mann-Whitney test: DN: Z=-0.696; p=0.514; DG: Z=-0.262; p=0.793 SG: Z=-1.049; p=0.316 CG: only small dogs' growl were used in this group), the size of the non-matching picture (DN: Z=-1.827; p=0.068; DG: Z=-0.145; p=0.887 SG: Z=-0.586; p=0.558 CG: Z=-0.614; p=0.551), nor the size (DN: Z=-0.176; p=0.86; DG: Z=-1.193; p=0.233 SG: Z=-1.182; p=0.237 CG: Z=-0.029; p=0.977) or the sex of the subject (DN: Z=-0.090; p=0.953; DG: Z=-0.088; p=0.930 SG: Z=-0.952; p=0.341 CG: Z=-1.026; p=0.305) had an effect on the looking preference.

Another interesting pattern was found when we compared the side preference in the looking behaviour of the dogs (Figure 30). In the cat-picture group a strong left gaze bias was found, while in the other three groups we found no such preference (Wilcoxon signed rank test: DN: p=0.1140; DG: p=0.5271; S: p=0.2897; **C: p=0.0087**).

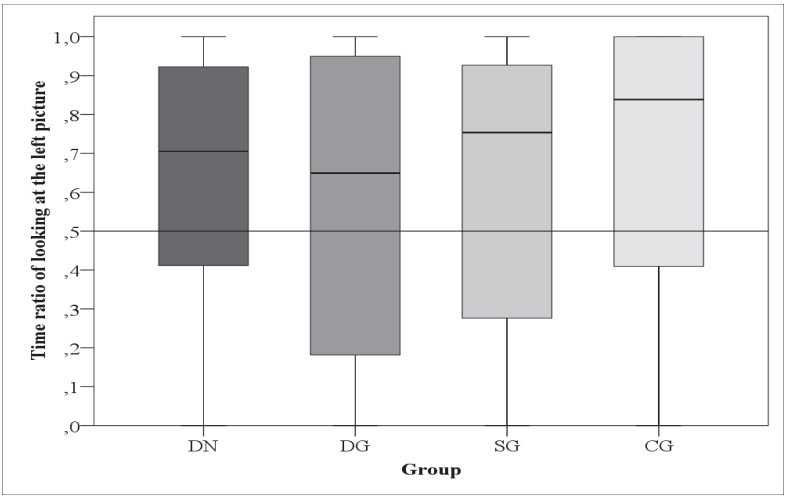

Figure 30. The side bias in the dogs' looking preference in the four experimental groups. The line at 0.5 time ratio indicates the lack of preference, the area above the line shows looking preference at the left picture. The boxes indicate the interquartiles, the whiskers the lowest and highest non-outlier values.

Discussion

The dogs' looking behaviour suggests that they can extract size information encoded in the growls of conspecifics. When confronted with two dog pictures while hearing the growl, dogs looked sooner and longer at the picture showing a dog matched in size to the growling dog, than a picture of a smaller or larger one, which means that they linked the acoustic size cue with the visual information provided by the pictures. However, our subjects did not generalize the size cue to any type of objects e.g., they showed no looking preference if the growl playbacks were paired with non-dog visual stimuli such as triangles or cat pictures. The fact that dogs had a looking preference exclusively when the two modalities were both showing dogs, suggests that our subjects linked the growls with the dog pictures, further suggesting that dogs might be able to activate a specific mental representation of the signaller with respect to the species and the size.

Riede and Fitch (1999) showed that the formant spacing in dog growls correlated not only with the vocal tract length, but also with the body size of the signaller. This suggests that formant dispersion can function as a reliable cue for dogs to assess the body size of conspecifics. Growls are mostly emitted as a threatening signal during agonistic social contexts, such as territory defence or food competition (Cohen & Fox 1976; Yeon 2007), thus the ability to estimate body size and fighting ability based on these vocalizations could

be advantageous for the receivers. As Taylor and co-workers (2008) have shown, formant dispersion is an important clue for humans to assess the size of dogs when they growl. In addition, they modelled in a playback study an intrusion of a strange dog to the subject dogs' household and they found that the subjects behave differently according to the perceived intruders' body size. Large dogs showed more explorative behaviours when they heard growls in which the formant dispersion showed smaller dogs than their own size (Taylor et al. 2010a). However, there was no other effect of the intruders' apparent body size when it was larger, or when the subjects were smaller. In contrast, our results showed that dogs are able to precisely match the size encoded in the growls with visual information about the emitter's size. Moreover, our findings strengthen the idea that looking preference in the test group was evoked by acoustic information provided by the growls. In contrast with Taylor and colleagues' results, we did not find any effect of the size of neither the growling dog, nor our subjects. This was probably caused by the difference between the experimental environments used during the growl playbacks. While Taylor's study modelled an intrusion into the household of the subjects, in which a more active explorative behaviour seems appropriate, our experiments were conducted in a strange place, where a more passive reaction can be expected from dogs.

In conclusion, we have shown for the first time that dogs can match cross-modal information between pictures and sounds and we provided evidence that dogs can assess accurately the size of a growling dog based on the acoustic information. Moreover, our results suggest that dogs are able to perceive species specific information based on pictures.

Experiment 4: Assessment of inner state and context recognition from growls by human listeners (Faragó, Takács, Pongrácz in prep)

Finally, we turn our focus from the role of growls in intraspecific communication to interspecific aspects. Several earlier studies suggest that asking humans to categorise animal calls by context or evaluate them with reference to emotional content provides a fruitful approach to find out how humans perceive animal vocalizations. There are two possible approaches in these studies. By playing back unknown animal sounds for humans to test their reactions we can extract whether similar or different mechanisms work during processing animal vocalizations. We can extract what acoustical parameters carry information for humans and we can find parallelities with animal communication and general rules how vocalizations can encode information (see Linnankoski et al. (1994) or Belin et al., (2008)). This approach can shed light on the evolutionary background of the vocal communication process. In the other hand, we can test how humans can use animal calls as source of information during communication, with providing vocalizations that are possibly used in joint social contexts with humans, like calls of domesticated animals (see Nicastro & Owren (2003) or Pongrácz et al. (2005)).

The first ones of such studies focused on perception of primate vocalizations. Brown et al. (1994) played back chirps of several East African birds (unidentified species) and other two monkey species (red tailed (*Cercopithecus ascanius*) and blue monkeys (*C. mitis*)) for humans and Sykes's monkeys (*C. albogularis*) in order to test and compare their discriminative ability by the means of psychoacoustic measures. They found that while the monkeys could discriminate more the monkey alarm chirps from non-alarm bird calls, humans differentiated all the three species categories' chirps (bird, red tailed and blue monkey). In another study human adults and children had to label different macaque (*Macaca arctoides*) vocalizations by its possible emotional background (Linnankoski et al. 1994). Humans were surprisingly successful. Children's performance improved with the age, and 9-10-year-olds did not differ from adults. The youngest group mostly misinterpreted the fear vocalizations and categorized them as angry or commanding. More recently, Owren & Rendall (2003) tested whether humans can discriminate rhesus monkeys' coos and screams, and showed that similarly to conspecifics, humans were better in discrimination of coo calls, although they performed above chance level in both call types. Humans also showed extremely good discriminative ability when they had to differentiate objects only by the playbacks of dolphin echolocation call reverberations (Au

& Martin 1989). Nicastro & Owren (2003) tested the functional referentiality of cat vocalisations by asking humans to classify them by their contexts. They found that the human listeners performed slightly above chance level, and the experience of the subjects with cats affected their context recognition. In contrast, experience with dogs did not affect the recognition of dog barks (Pongrácz et al. 2005), nor the association between barks and affective states (Pongrácz et al. 2006). Moreover the fact that blind people without any previous visual experience were as successful as sighted humans supports that the dog barks carry contextual information in their acoustic structure (Molnár et al. 2009a). Interestingly humans were not good in individual discrimination (Molnár et al. 2006). In contrary the common sense owners could not differentiate their own dog's barks from other individuals of the same breed. Besides their poor performance, the tonality of the barks affected the humans' individual differentiation; they could differentiate noisier dog barks recorded from agonistic contexts.

Human perception of acoustic cues in dog growls were already tested in some studies by Taylor et al (2008). They found that both fundamental frequency and formant dispersion conveys size information for humans: dogs emitting growls with lower F0 and closer formats were rated larger (Taylor et al. 2008). Human listeners rated such growls also as more aggressive (Taylor et al. 2010b). Similar pattern occurred when they tested humans with play and aggressive growls and they found that humans could not differentiate the context of single growls (Taylor et al. 2009). Moreover, they found that using resynthesized bouts of growls improved their performance. Where the inter-growl intervals were manipulated to refer typical growling rates in play or aggression contexts, humans categorized them correctly above chance level.

In our study we played back natural growl sequences for human listeners without manipulating their frequency and time structure. While Taylor et al (2009) aimed to test how acoustical parameters affect the recognition of dog growls, we focused more on how humans are able to convey emotions and recognize the contexts of unmodified, natural growls. We used three contexts, and added the food guarding growl to the playful and threatening contexts used by Taylor et al. (2009). We asked our subjects to rate the inner state of the dogs not just on the playfulness and aggressiveness scale but on three other emotional scales, see Pongrácz et al. (2005). We assumed that humans will perform better than reported by Taylor et al (2009), and recognise context and supposed inner state of the dogs by natural growl sequences, as Nicastro & Owren (2003) found in cat meows and Pongrácz et al. (2005) in dog barks.

Materials and methods

Subjects

40 adult human subjects participated in our experiment (14 male and 26 female, for details see Appendix Table 9). Participation was voluntary, and the subjects had no prior information about the questions and goals of the study. Subjects were tested one by one, with the presence of the experimenter (N. T. or T. F.).

Stimuli

For the playbacks we used our dog growl recordings from the pool of vocalization sequences collected for the acoustical analysis. Three contexts were represented in this study: guarding food from a conspecific (FG), threatened by a strange human (TS) and playing tug-of-war with the owner (PL). We cut 10 seconds long sections from the original recordings, and we choose those sequences for playback which contained at least three growls with low background noise. This way we got eight different growl samples in each contexts recorded from 18 different dogs (Appendix Table 10). Then we generated in a semi-random way twenty different sequences containing six growl samples, two-two from each context. In one sequence there were no two consecutive growls from the same context (Appendix Table 11). During the playbacks two subjects heard the exact same sequence, and ten subjects judged the exact same growl sample.

Set-up

The tests were performed in a silent room (Video lab) of the Department of Ethology. During the playback session only one subject and one experimenter were present in the room. After a short explanation of their task the subjects got the first questionnaire in which they had to score the emotional content of first six growl samples. Each growl was played in a sequence after a half minute long pause, and the subjects could easily rate the growls one by one. After the end of the first sequence we gave the subject the second questionnaire for categorizing the context of the growls in the second sequence. We played the second sequence in a similar way than the first one. If the subject had difficulties with one growl, we played it again for a second time. At the end of both sequences we repeated the first growl of the sequence for measuring the reliability in the response of the subjects.

Questionnaires

In the beginning of the questionnaire we asked the subjects for some basic background information on age, experience with dogs, and being bitten by a dog. For the emotional scaling the subjects had to rate the growls by five inner states: aggression, fear, despair, happiness, playfulness (see (Pongrácz et al. 2005; Pongrácz et al. 2006). However, in contrast to these studies we used a visual analogue scale (VAS) for scoring. Subjects had to place a mark on a 10 centimetres long line. The distance of the mark from the left end of the line in millimetres represented how much they felt that given inner state related to the actual growl. This way the subjects had bigger freedom in the scaling and this method provides a finer discrimination of the growls and also results in a continuous variable for our measurements (Maxwell 1978).

On the context questionnaire the subjects had to choose one from the three possible contexts. The subjects filled in the emotional questionnaire first because we tried to avoid the impression that there are only three possible contexts (for scoring the emotional content).

Data analysis

We measured the scaling of the different emotions with a ruler. The data were recorded in millimetres growl by growl for each inner state. The responses only on the first six growl samples were used for both analyses. The averages of the scaling were calculated for each inner state. For the context questionnaire we recorded the responses of the subjects as a nominal variable, and compared it with the actual context of the growls. The correct answers were counted. For the confusion matrix the percentage of the three given answers for all three growl types were calculated. We also calculated the average scaling of each inner state growl by growl for correlation studies.

Statistics

Nonparametric tests were used because our data did not follow the Gaussian distribution. We compared the scaling of the subjects between the three types of growl by the five inner-states with Kruskal-Wallis test, and within the growl types we collated the emotional scaling with Friedman test, for post hoc comparison Dunn test were used.

To test the possible effect of the subject background variables (age, experience with dogs, dog bite history) we averaged the scalings of the same context growls within subjects, thus we got one data point in each context per individual. Similar pooling was applied to test

the effect of known acoustical (fundamental frequency, formant dispersion and growl length) parameters and the size of the growling dog on the emotional scalings, but now we pooled each rating of the same growl sample.

All the correlations were analysed with Spearman test including the consistency of the emotional scaling and the context recognitions. Binomial tests were used to compare the subjects' context choice to the random 33% correct classification. Due to the multiple tests FDR correction was applied to avoid false discovery.

Results

Assessment of emotions in the growls

Both within and between the growl types marked significant differences were found (Figure 31). The food guarding growls were scored as being highly aggressive, less fearful and desperate and the happiness and playfulness scores were significantly lower than the other three negative inner states (Friedman test: $\chi2(4)=105.59$; $p<0.001$). The growls emitted at threatening humans were found also aggressive and more desperate and fearful than playful or happy (Friedman test: $\chi2(4)=55.87$; $p<0.001$), while play growls showed just the opposite pattern: they sound highly playful and happy for the human listeners, and were given low scores on the aggression, despair, and fear scales (Friedman test: $\chi2(4)=83.93$; $p<0.001$). Between the growl types all the five scalings were different (Kruskal-Wallis tests – playfulness: $\chi2(2)=82.61$; $p<0.001$; fear: $\chi2(2)=38.41$; $p<0.001$; aggressiveness: $\chi2(2)=43.01$; $p<0.001$; happiness: $\chi2(2)=62.45$; $p<0.001$; desperate: $\chi2(2)=18.75$; $p<0.001$). The play growls differed significantly in all the inner states: they have been scored more playful and happy than the two agonistic growls and less desperate, fearful and also lowly aggressive. We found no significant differences between the scaling of the food guarding and threatening stranger growls except and most importantly in the aggression motivational state: subjects found that the threatening stranger growls sounded less aggressive.

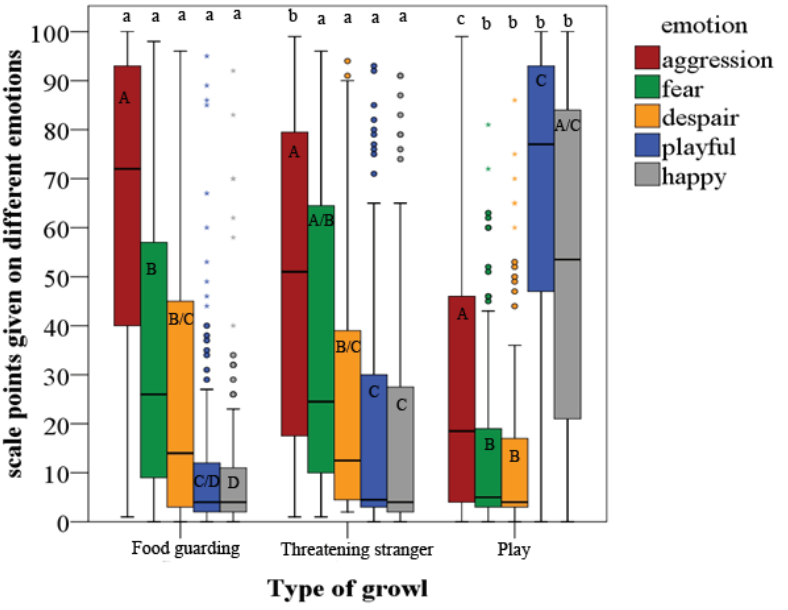

Figure 31. Medians of the emotional ratings between the three growl types. The boxes indicate the interquartiles, the whiskers the lowest and highest non-outlier values, while the circles and stars the outliers. The different capital letters show significant differences between the emotional scalings within contexts. The letters on the upper side of the graph indicate significant differences between contexts. Both results obtained from the Dunn post-hoc tests.

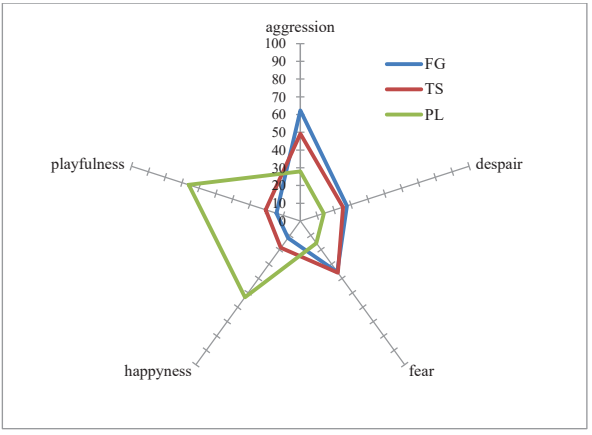

Figure 32. Spider web graph of the average emotional scalings. The areas represent the emotional state space derived from the humans' subjective ratings.

The Spearman tests checking the consistency of the subjects' response showed that only the scoring of the fearfulness did not correlate significantly. In the case of the other four inner states scaling correlated positively between the first and second time (Spearman correlation tests – playfulness: δ=0.63; p<0.001; fear: δ=.167; p=0.303; aggressiveness: δ=0.59; p<0.001; happiness: δ=0.62; p<0.001; desperate: δ=0.38; p=0.015).

Independently from the contexts there were no effect of the subject gender, experience with dogs or bite history (Mann-Whitney tests see Table 4.). Age of the subjects had no significant effect on the ratings either.

Food guarding						
		play	fear	aggression	happiness	despair
gender	U	159.000	159.000	158.500	179.000	177.500
	p	.514	.514	.505	.932	.898
experience	U	132.000	103.000	125.000	95.000	95.000
	p	.808	.237	.639	.149	.149
bitten	U	172.500	171.000	180.000	179.000	154.500
	p	.590	.562	.740	.719	.300
Threatening Stranger						
gender	U	103.000	108.000	113.000	106.000	170.500
	p	.025	.036	.050	.031	.744
experience	U	103.000	129.500	86.000	100.000	136.000
	p	.237	.746	.083	.200	.910
bitten	U	162.000	172.500	174.500	166.500	191.500
	p	.407	.590	.629	.481	.989
Play						
gender	U	177.000	165.500	151.000	170.000	148.500
	p	.887	.639	.379	.734	.341
experience	U	99.500	99.000	127.000	91.500	89.000
	p	.195	.189	.685	.120	.101
bitten	U	175.000	179.500	191.000	177.000	178.000
	p	.639	.730	.978	.679	.699

Table 4. Effect of the subjects background variables on the emotional scalings of the growl samples. The significance level was corrected to 0.01 with FDR.

We also found no correlation with the acoustical parameters in the frequency domain but one occasion, fundamental frequency showed weak positive correlation with playfulness scores (ρ(24)=0.437; p=0.033). In the time domain the length of the individual growls correlated with all inner state scaling: positively with fear (ρ(24)=0.769; p=<0.001), aggression (ρ(24)=0.668; p=<0.001) and despair (ρ(24)=0.523; p=0.009), and negatively with happiness (ρ(24)=-0.803; p<0.001) and playfulness (ρ(24)=-0.77; p<0.001).

Context recognition

Overall subjects classified correctly 63 % of the growl samples that is significantly higher than the 33 % chance level (Binomial test: $p<0.001$). All the three growl types were recognized also above chance level. The human listeners classified almost 60 % of the food guarding growls correctly, also half of the threatening stranger and 81 % of the play growls were recognized. Present results show that subjects distinguished play growls easier compared to the two agonistic ones and the confusion matrix showed that a relatively high amount of threatening stranger growls were considered to be food guarding and vice-versa (Figure 33).

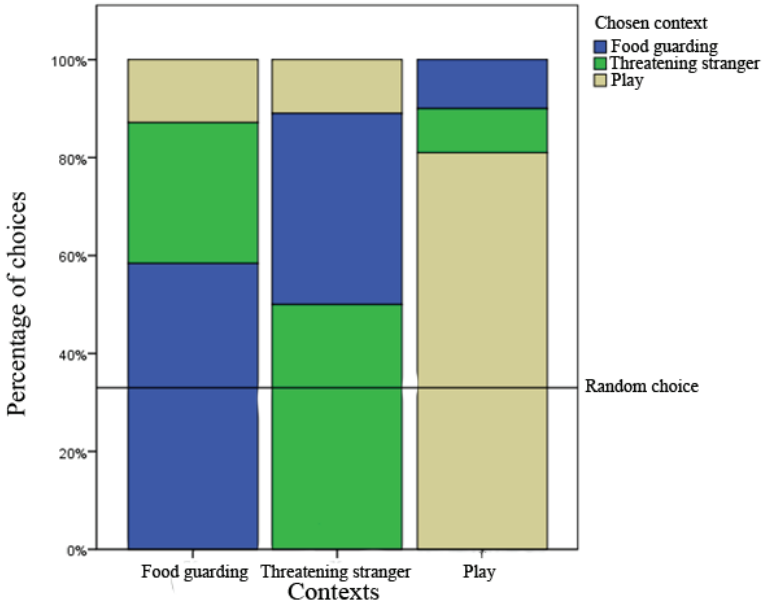

Figure 33. Distribution of context choices of the subjects. The confusion of the two agonistic growls is clearly visible.

Similar pattern is discernible in the comparison of the responses given to the first and the seventh playback. We found that the subjects' choices correlated positively between the two identical growl samples (Spearman: $\delta=0.612$, $p<0.001$). However, their choices were only roughly consistent (Figure 34), mainly because the discrimination of the two agonistic growls was uncertain. Interestingly, the second choices were more accurate: while only 19

subjects classified the growls correctly after the first hearing, for the second round 26 subjects chose the correct context.

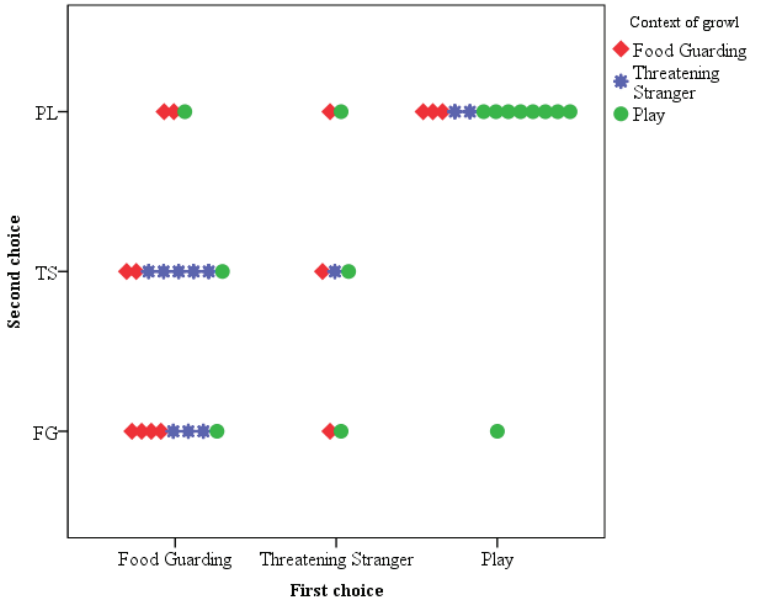

Figure 34. The consistency of the context choices between the first (horizontal axis) and seventh (vertical axis) playback.

Similarly to the inner state ratings, the subjects' age, gender, experience with dogs and bite history did not affect their success in guessing the context of the growls. The subjects' success in guessing the correct context of the growl samples did not correlate with any of the acoustic or size parameters.

Discussion

We found that humans were able to recognise the context of growls produced by dogs in various situations, although their performance was lower with agonistic growls. Their error pattern clearly showed that most of the wrong choices were intermixes of the food guarding and threatening stranger situations. The inner-state scaling results reflected that our subjects could attribute adequate emotions to growls. They rated the growls emitted during play high on the happiness and the playfulness scales, while the fear and despair ratings were low in all three contexts. Both agonistic growls were rated high on aggression, but the subjects judged the food guarding growls more aggressive. This latter finding is

especially interesting in the light of our earlier studies in which we could not find differences in the measured acoustical parameters, but the dogs reacted differently to the two agonistic growls (Faragó et al. 2010b). This shows that both humans and dogs can sense the difference between these two contexts based only on the acoustic information.

We found also that mainly the length of the growls and the pulsing of the growl sequence affected humans' rating. This finding is in accordance with the results of Taylor et al. (2009) who showed that humans could categorize and attribute correct inner state to the resynthesized growl bouts in which the temporal structure was similar as in natural growls. Nicastro & Owren (2003) showed similar ease in context recognition in cat vocalizations. Rhythm and inter-bark intervals were considered as an important cue for humans to assess the dogs inner state (Pongrácz et al. 2006): longer pauses between individual barks were judged as indicating lower aggressive tendencies. The opposite is true in the case of growls. The play growls are usually characterized by fast growl sequences with short intervals (Taylor et al. 2009). An important difference to the Taylor's study is that the growl sequences used contained uniformly long growls. However, natural growls differ in length that provides an additional feature for humans for being able to deduce the context.

In our study nor the emotional ratings, nor the context recognition were affected by the dogs' size cues as reported by Taylor et al (2010a). The exception was that our subjects found higher pitched (fundamental frequency) growls more playful. This discrepancy could be explained by the fact that Taylor at al. utilized resynthesized individual growls for which the formants and fundamental frequency were manipulated in order to indicate larger or smaller individuals. This may have resulted in clearer size cues for the listener while in our natural growls this information may have been overshadowed by other acoustic features. The rating of higher pitched growls as more playful supports Morton's motivation-structural rule (Morton 1977).

General discussion

In our studies we found clear evidence that growls convey information about the accurate size of the signaller, its inner state and the context of the vocalization. We found also that both human and conspecific listeners are able to process and rely on such information contained in the growls.

In the first study we revealed differences in the acoustic structure of between growls recorded in agonistic and playful contexts. These differences appeared in both time and frequency domain: playful growls were shorter, had higher fundamental frequency and lower formant dispersion. Taylor et al found similar pattern in the time domain upon comparing playful and aggressive growls (Taylor et al. 2009). In contrast to our findings, they found no difference in the fundamental frequency or in the formant dispersion. This contradiction may be due to the different recording circumstances, although in that study there is no clear description of the playful context: "the owner encouraged their dog to play, usually by showing it a toy as stimulation and giving vocal encouragement" (Taylor et al. 2009). Thus it is possible that the actual playful context differed in the two studies. Our subjects growled during a tug-of-war game, while it may be that Taylor's subjects played with the owner in a different way.

Play is paradoxically not what it appears to be (Bateson 1955). During social play, animals behave ambiguously: they show signals of aggression and submission, they mock attacks and bites, and then retreat and run away (Sutton-Smith 2001). In the course of the play they may alternate roles (Fagen 1981). During this they show playful communication signals. According to Bekoff and Allen this communication can happen in two ways: by the recognition of subtle differences in behaviour or with contextual cues that inform the partner about playfulness (Bekoff & Allen 1998). The former occurs for example in coyote play during which the players direct their bites in play fight towards different places of the body then during real fights (Hill & Bekoff 1977). In the latter case play signals, such as the play-bow in canids (Bekoff 1995), set the context of playing. Upon encountering a threat signals followed by a play bow, coyotes will react with play. However, if the play-bow is absent the threat will interpreted as a real threat (Bekoff 1975). These play signals decrease the misinterpretation of the aggressive signals and prevent the escalation of real fights (Pellis & Pellis 1996).

In the light of these findings, our results on play growls can be interpreted as play signalling. The lower formant dispersion communicating bigger body size can be part of

the playful repertoire as exaggeration (Fagen 1981), while the higher fundamental frequency might communicate the lack of aggression and playfulness. Acoustic play signal have been described in the literature, for example, bottlenose dolphins use a dedicated pulse-whistle during play fights that is similar to the aggressive pulses, but shorter, and pulsing faster (Blomqvist et al. 2005). In primates the visual signals like the play face in chimpanzees and bonobos (*Pan paniscus*) (Flack et al. 2004; Palagi 2008), head shaking in spider monkeys (Pellis & Pellis 2011) or tail signals in lemurs (Palagi 2009) are also accompanied by play vocalizations that thought to be the antecedent of the human laughter (Vettin & Todt 2005). These play pants in chimpanzees, however, seems to serve as signal for sustaining play interaction signal because they are rarely emitted by the playful aggressor (Matsusaka 2004). In contrast, in Barbary macaques Kipper & Todt (2002) found that the pants supplement the play face and do not substitute it. We can also interpret play growls during tug-of-war games as a sustaining signal if we notice that if the dog would like to get full possession of the tug it should just stop growling and attack the other, instead of pulling and growling vigorously. Play growls could also communicate the excitement and willingness to participate in and go on with the game. However, there is little evidence whether any aspect of the acoustical structure of the play growls would support it acts as a play signal. It is also not clear whether these play growls can considered as signalling active dishonest (similarly to the size exaggeration attempt of the roaring red deer), or the acoustic modification is just a consequence of the mouth and head position during the play. More detailed analysis would be necessary to decide if active laryngeal movements and retraction are involved in this vocalization as Fitch & Reby (2001) revealed it in the case of barks. Also recording and analysing growls emitted in other playful contexts (like wrestling with other dogs) would provide good additional data on this question.

In our third experiment we found that dogs are able to assess accurately the size of a growling individual. As already mentioned several species show similar ability of spontaneous perception of size information providing indirect evidence on size assessment abilities, however we know about only a handful attempt to test the animals' ability to use this cue to assess the physical attributes of conspecifics. In red deer clear association was shown between body size and formant dispersion (Reby & McComb 2003b). However, all the previously mentioned playback studies provide behavioural evident on sensitivity of this size cue but still no assessment. Also in some birds the is passive size exaggeration by the elongation of the trachea (Fitch 1999) suggests that these species might possess the

ability to use formant dispersion as size cue. Whooping cranes (*Grus americana*), for example, showed dishabituation after hearing calls with modified formant dispersion (Fitch & Kelley 2001). In this study, some subjects showed approach reaction to the source of the calls indicating smaller individuals. In the light of that adult cranes show a general interest towards young individuals, this can suggest that they could assess the size of the modelled caller and expect the callers to be youngsters. Apart from extensive work on humans, only rhesus macaques were tested and provided data supporting the presence of size assessment ability in animals (Ghazanfar et al. 2007) besides our result on dogs. According to Taylor et al. (2008) and Riede & Fitch (1999) agonistic growls seem to act as honest signals, at least with regard to the size information encoded in the formant structure. Our results support also these findings, and Taylor et al reported a further replication of these results, after our report was published (Taylor et al. 2011). There was some difference in the method, however. While we used video projections and adjusted the size of the dog pictures to the size of the growling dogs, Taylor et al used 3D dog models: they showed a large stuffed German shepherd and a small Jack Russell terrier which were presented with resynthesized growls to match their size. Their results support our findings: the dogs preferred to look at the matching model indicating that they are able to use the formant dispersion as a size cue. However Taylor's study was limited only to two 3D models in showing that dogs are spontaneously sensitive to the size information and still shows no evidence on size assessment. In contrast, our study provided evidence for a larger scale of different size categories, suggesting accurate size recognition and we gave deeper insight of how dogs can use this size information and how they match visual and acoustical information.

Our study provides some insight to the two dimensional visual perception of the dogs too. Although we have little knowledge about how animals perceive the two dimensional representations of real world objects, numerous studies on animals and humans use pictures instead of real objects to investigate cognitive abilities (Bovet & Vauclair 2000). Several studies have used pictures to test individual and species specific recognition (for example in sheep (Bouissou et al. 1996) and cattle (*Bos primigenius taurus*) (Coulon et al. 2007)). There are some studies which suggest that dogs are also capable of extracting some information from pictures. Young and adult dogs sniffed those areas of a life-size dog painting which are normally investigated by dogs during an encounter with a strange conspecifics (Fox 1971b). Investigating dogs' performance in a pointing and an obedience task, and Pongrácz et al. (2003) found that the dogs readily obeyed the command of the

owner if a life sized visual image of the owners was projected on a the wall (Pongrácz et al. 2003). Dogs can also discriminate dogs and human pictures based on facial cues (Racca et al. 2009), moreover Range and co-workers showed by the means of a touch-screen method that dogs were able to classify photographs of dogs and landscapes correctly based on perceptional differences like class specific contrast, luminance or colour features (Range et al. 2008). Fagot and co-workers differentiated three ways of pictorial processing: Independence, Confusion and Equivalence (Fagot et al. 2000). In independence, the subject makes no connection between the picture and its referent treating them as independent, different entities besides their featural similarities. While equivalence means that the animal perceived the picture as a representation of a real object. This says that the picture acts as an iconic, symbolic representation of the referred object and this assumes that the subject has high cognitive abilities. In the third case, the viewer of the picture confuses it with its content and acts like it would be the real three-dimensional object (Fagot et al. 2010). We assume that the dogs in our study worked on this third case in order to link the visual and acoustical modalities. The dogs in our study showed significantly more attention towards the pictures if the pictures showed dogs or cats. Moreover, the fact that the dogs showed a preference towards the matching sized picture only on the occasion of dog-growl intermodal pairings, suggests that they could perceive the dog pictures as dogs. However to confirm this, other, more specific tests would be necessary, such as showing pictures of different species pairs, such as dogs versus cats, while playing dog and cat specific calls, and additionally an unknown control vocalization. In such experiment we can assume that the dogs can recognize (confuse) dog pictures, they will show clear looking preference towards them when hearing dog specific calls, while the cat vocalization should evoke preference towards the cat picture, and the unknown vocalizations should cause no preference.

In an earlier study, Adachi and co-workers demonstrated that dogs looked longer (assumed surprise reaction) if a picture of a stranger's face coupled with the voice of their owner was shown to them, and similarly they looked longer at the owner's picture when a stranger's voice was played back (Adachi et al. 2007). Such intermodal matching ability has been shown in several other species recently in connection with various different cognitive abilities (individual recognition in a chimpanzee: (Martinez & Matsuzawa 2009); recognition of communicative signals in capuchin monkeys (*Cebus capucinus*): (Evans et al. 2005) facial expressions in rhesus monkeys: (Ghazanfar & Logothetis 2003) and size perception: (Ghazanfar et al. 2007)). Our results and the aforementioned studies, suggest

that dogs as well as primates, have the ability to match the mental representation of the signaller with its vocalization. In our study, the overall duration dogs spent observing the pictures did not vary if they heard a growl or a noise and also did not differ between the groups that saw dog or cat pictures. However, the triangles grabbed the attention of the dogs to markedly lesser extent. Complex pictures of conspecifics' face might contain more processable information for dogs than triangles, and key stimuli, like the eyes of the face, may elicit enhanced attention.

We also realize that it is possible that the longer looking times at the dog and cat pictures can be attributed to different causes. Looking at dog pictures can be an attentive response, elicited by perceiving conspecifics, while the cat picture in odd pairing with dog growls, can cause a surprise response with orientation behaviour. This is supported by the fact that dogs showed marked left gaze bias only when looking at the cat pictures. This interesting result can be explained by the asymmetric manner of perception of sensory stimuli. Recent studies showed a left gaze bias in dogs in the case of simultaneous bilateral heterospecific visual presentations (cat and snake silhouettes), (Siniscalchi et al. 2010) as well as acoustical (Siniscalchi et al. 2008) unanimated natural sound stimuli (thunder). More importantly, when dogs were presented with a dog silhouette or dog vocalizations, they showed a right gaze bias or no bias et all in the above mentioned experiments, suggesting that similarly as in other mammals including humans, the right hemisphere shows increased activity in association with novel, attention grabbing or unexpected stimuli (Rogers & Andrew 2002). Because the auditory tract fully crosses in the dogs' brain and the optical nerve crosses 75% as well (Fogle 1992), the left gaze bias towards cat pictures, coupled with growls found in our study, could be explained by the non-matching visual and acoustical modalities functioning as a novel, non-expected stimulus that is processed more by the right hemisphere causing a left gaze bias. The lack of a right gaze bias in our study in contrast to findings of Siniscalchi and co-workers' experiments can be explained on the basis of methodological differences. In our case the sound was played back form a loudspeaker positioned between the two picture presentations while both Siniscalchi et al. (2008) and Siniscalchi et al. (2010) utilized bilateral presentation of dog vocalisations evoking stronger side effect. It is also possible that our experimental manipulation masked out this gaze effect because the size information might have had a stronger impact on the looking behaviour of the dogs in cases when the matching picture was presented on the left side.

Although we did not find measurable acoustical difference between the threatening growl and food-guarding growl both dogs and humans seem to sense differences between these two types of agonistic vocalisations. The dogs showed stronger withdrawal reaction to the food guarding growls in the appropriate context, while the threatening stranger growls did not deter dogs to approach and take away or eat the bone. Humans also rated the threatening stranger growls significantly less aggressive (however still they judged them as being more aggressive than playful growls) and they were able to link the two agonistic growls to their appropriate context. This suggests that growls may convey contextual information, in addition to having possibly graded nature. Similar signals are known in the literature on functionally referential communication: distinct variants of graded signals can be used in different contexts, and fall in different perceptual categories in the listeners. Baboon grunts (Owren et al. 1997), wahoo-s (Kitchen et al. 2003) and barks (Fischer et al. 2001a) all have variants that vary on a continuous acoustical scale, but are distinct enough for clear discrimination and used in different social contexts (Rendall et al. 1999; Fischer et al. 2001b). In the case of baboon barks noisiness is the main graded variable: tonal barks used as contact calls, emitted by isolated individuals, while noisy barks function as alarm signals. Interestingly in a habituation-dishabituation test, baboons failed to discriminate the less noisy, intermediate barks from the contact calls suggesting that the noisiest variants are categorized by the subjects into a distinct bark type (Fischer et al. 2001b). As the occurrence of nonlinear events such as deterministic chaos in the calls increases with the arousal of the individual, the more noisy calls refer to higher urgency as in suricatas (Townsend & Manser 2011) and marmots (Biedenweg et al. 2011) and possibly in the case of baboons, where barks are more associated with the presence of a predator. The alarm barks of baboons can be further classified in two main subtypes given on the proximity of lion or crocodile based on the level of the fundamental frequency modulation (Fischer et al. 2001a).

Recently it was shown that several chimpanzee have context dependent variants like the barks (Crockford & Boesch 2003) the rough grunts (Slocombe & Zuberbühler 2005a) pant hoots (Notman & Rendall 2005) and agonistic screams (Slocombe & Zuberbühler 2005b). The agonistic screams are not just context specific, but convey information about the social role (aggressor or victim) of the caller during the fight, and listeners look longer if presented with incongruent playbacks pretending a low ranking individual as an aggressor above a high ranking one (Slocombe et al. 2010). However, in this case we cannot talk about functional referentiality of the calls because the aggressor and the victim

are in a different inner state losing one criteria of a referential call system. We can say that these calls provide only context specific information, listeners probably link both the aggressor's and the victim's call with the same context.

We should note that the conceptual problem of the unknown inner state of the callers could apply to many earlier studies on functionally referential food and social calls. For example, compare the present results with the study by Slocombe and Zuberbühler (2005) which presents the most similar behavioural situation to the ours. In both cases signallers (dogs or chimpanzees) emit vocalizations in social contexts (dogs: defensive aggression; chimpanzees: aggression/ defence in an agonistic scenario). If we examine carefully the situations in both cases, we could argue that it is more likely that dogs in the two agonistic contexts have more similar inner states than the chimpanzees, in which it is more likely that the 'aggressor' chimpanzee is in different inner state than the 'victim' (see the introduction of Experiment 2). Furthermore, social referents, such as "victim" or "aggressor", or in our study the dogs that show defensive aggression or food guarding are maybe less clear-cut examples of functionally referential communication (compared to "predator class" for example). It is difficult to know in these cases how "external" the referent really is.

In conclusion, our studies provided new insights in the vocal communication of dogs. Our results showed the diverse communicative capacities of a relatively simple vocalization, and shed light how the growls carry information about the context, and the physical features of the signaller for conspecifics and humans.

Outlook

However, several further questions remained open. For example, it is necessary to collect more growl samples from as many individuals and from the same subjects in as many contexts as possible for more detailed acoustical analysis. If we have a high number of growls, we could apply Principal Component Analyses to extract those parameters that may carry contextual information, or use permuted Discriminant Function Analysis for finding the acoustical difference between the two types of agonistic growls.

The development of growl use is also a question for further studies. We know that growls appear in the pups' repertoire early during the ontogeny (Bleicher 1963). A longitudinal collection of growl samples from an early age could reveal the acoustical divergence of the agonistic and playful growls. Moreover, with similar playback studies as

the second experiment, we would be able to test the role of learning in the contextual differentiation of the agonistic growls in puppies.

For studying the interesting ambiguity of size coding in play growls an experiment is already running in the Department of Ethology. Using a slightly modified method of the looking preference paradigm we playback playful growls in parallel with differently sized dog pictures to test whether dogs use the size information for assessing the other dogs' size.

A further interesting question is how domestication and directional selection of hunting or fighting dogs altered their ability of size assessment. We may hypothesise that in the case of those breeds (e.g. like small fox hunting terriers), which were selected for fighting with other animals that are far larger than themselves, may have an impaired ability to assess the strength of others based on vocal cues. Thus such breeds may not show preference in a crossmodal playback study using agonistic growls.

For studying the gradedness of the growls, and how dogs potentially alter the acoustical parameters to communicate different levels of inner states we can use modified recording contexts in which we vary the evoking stimuli. For example, in the threatening stranger context we can use male vs. female human for threating the dog, or humans with different build. In the food guarding context we could use different sized dogs as strangers and compare the acoustical structure of these growls and extract the parameters that are possibly modified during vocalization according the different level of the inner state of the dogs. These growls could be a good source for another playback study with humans to test how their subjective emotional scaling changes with the different evoking stimuli and the changed acoustical parameters.

Finally we conclude that the diverse vocal repertoire of dogs makes them a valuable source of studying general questions of animal communication such as the functional referential signalling. Our results on growls have provided exciting insights into the communication of our long lasting companion, the family dog.

Figure 35. Recording or playback of wolf vocalizations in a zoo. (Taken from: http://blackandwtf.tumblr.com/post/434836343/via-emma)

Acknowledgements

I would like to thank to the Hungarian National Science Foundation (No. T047235 and No. K58020) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences for providing the financial funding. Our studies were also funded by a grant from the European Union (FP7-ICT-2007 LIREC 215554) and the ETOCOM Project (TÁMOP-4.2.2-08/1/KMR-2008-0007) through the Hungarian National Development Agency in the framework of Social Renewal Operative Programme supported by EU and co-financed by the European Social Fund. My guest research scholarship at the Clever Dog Lab (Department of Cognitive Biology, University of Vienna) was funded by the Aktion Österreich-Ungarn Foundation (No. 68öu4 and No. 74öu3) and the Royal Canin.

First of all, I am thankful of Dr. Péter Pongrácz, my supervisor for his support, kind leadership, fatherly advices, bright ideas and because I can call him as my friend, of Dr. Ádám Miklósi the Professor of the Department of Ethology for providing the opportunity to work at the Department, for his objective view on our work and his advices that pushed our studies forward. I am grateful for the technical support and everyday companionship of Antal Dóka, without him we would be definitely stuck. I also appreciate the statistical advices of Dr. András Kosztolányi and all the help that were provided by the other colleagues at the Department. Also my work would be nowhere without the help and hospitality of Dr. Zsófia Virányi, Dr. Friederike Range, Dr. Ludwig Huber and all the colleagues of the Department of Cognitive Biology and the Clever Dog Lab in Vienna. I am grateful of Celeste Pongrácz for correcting the English in my manuscripts, of Thierry Aubin and several anonymous reviewers for their helpful criticism.

I wish to thank my friends Nándor Takács and András Péter the work, the talks and the laughs together in the Video Room and at several other nice places.

The kindest thanks are entitled to Zsuzsanna Szatmári, my fiancée for her love, support and belief in me, for her gentle enforcement to finish my thesis and avoid procrastination. I am also grateful of my family, my mother and sister for supporting me through all the years.

At least but not last, I would like to thank all the owners and dogs in Budapest and Vienna for their participation in my studies.

References

Adachi, I., Kuwahata, H., & Fujita, K. 2007. Dogs recall their owner's face upon hearing the owner's voice. *Animal cognition*, **10**, 17-21.

Adams, E. S., & Caldwell, R. L. 1990. Deceptive communication in asymmetric fights of the stomatopod crustacean Gonodactylus bredini. *Animal Behaviour*, **39**, 706-716.

Arch, V. S., Grafe, T. U., Gridi-Papp, M., & Narins, P. M. 2009. Pure ultrasonic communication in an endemic Bornean frog. *PloS one*, **4**, e5413.

Arnold, K., & Zuberbühler, K. 2006. The alarm-calling system of adult male putty-nosed monkeys, Cercopithecus nictitans martini. *Animal Behaviour*, **72**, 643-653.

Arnott, G., & Elwood, R. W. 2009. Assessment of fighting ability in animal contests. *Animal Behaviour*, **77**, 991-1004.

Au, W. W. L., & Martin, D. W. 1989. Insights into dolphin sonar discrimination capabilities from human listening experiments. *The Journal of the Acoustical Society of America*, **86**, 1662-70.

Aubin, T., & Jouventin, P. 1998. Cocktail-party effect in king penguin colonies. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 1665.

Bahrick, L. E., & Hollich, G. 2008. Intermodal perception. *Encyclopedia of Infant and Early Childhood Development*, **2**, 164-176.

Bardeleben, C., Moore, R. L., & Wayne, R. K. 2005. A molecular phylogeny of the Canidae based on six nuclear loci. *Molecular phylogenetics and evolution*, **37**, 815-31.

Barnard, C. J. 2004. *Animal behaviour: mechanism, development, function and evolution.* Harlow: Prentice Hall.

Bateson, G. 1955. A theory of play and fantasy. *Psychiatric Research Reports*, **2**, 39-51.

Beil, R. G. 1962. Frequency Analysis of Vowels Produced in a Helium-Rich Atmosphere. *The Journal of the Acoustical Society of America*, **34**, 347.

Bekoff, M. 1975. The communication of play intention: Are play signals functional? *Semiotica*, **15**, 231-240.

Bekoff, M. 1995. Play signals as punctuation: The structure of social play in canids. *Behaviour*, **132**, 419–429.

Bekoff, M., & Allen, C. 1998. Intentional communication and social play: how and why animals negotiate and agree to play. In: *Animal Play*, (Ed. by M. Bekoff & R. D. Byers), pp. 97-114. Cambridge, UK: Cambridge University Press.

Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M. D., & Armony, J. L. 2008. Human cerebral response to animal affective vocalizations. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 473-81.

Biedenweg, T. A., Parsons, M. H., Fleming, P. a., & Blumstein, D. T. 2011. Sounds Scary? Lack of Habituation following the Presentation of Novel Sounds. *PLoS ONE*, **6**, e14549.

Bleicher, N. 1963. Physical and behavioral analysis of dog vocalizations. *American Journal of Veterinary Research*, **24**, 415-427.

Blomqvist, C., Mello, I., & Amundin, M. 2005. An Acoustic Play-Fight Signal in Bottlenose Dolphins (Tursiops truncatus) in Human Care. *Aquatic Mammals*, **31**, 187-194.

Blumstein, D. T. 1999a. The evolution of functionally referential alarm communication. *Evolution of Communication*, **3**, 135-147.

Blumstein, D. T. 1999b. Alarm calling in three species of marmots. *Behaviour*, **136**, 731-757.

Blumstein, D. T., & Armitage, K. B. 1997. Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Animal Behaviour*, **53**, 143-171.

Boersma, P., & Weenink, D. 2001. Praat, a system for doing phonetics by computer. *Glot International*, **5**, 341–345.

Boitani, L., & Ciucci, P. 1995. Comparative social ecology of feral dogs and wolves. *Ethology Ecology & Evolution*, **7**, 49–72.

Bouissou, M. F., Porter, R. H., & Ferreira, B. G. 1996. Influence of a conspecific image of own vs. different breed on fear reactions of ewes. *Behavioural Processes*, **38**, 37-44.

Bovet, D., & Vauclair, J. 2000. Picture recognition in animals and humans. *Behavioural brain research*, **109**, 143–166.

Bradbury, J. W., & Vehrencamp, S. L. 1998. *Principles of animal communication.* Sunderland, Massachusetts: Sinauer Associates.

Brown, C. H., Sinnott, J. M., & Kressley, R. a. 1994. Perception of chirps by Sykes's monkeys (Cercopithecus albogularis) and humans (Homo sapiens). *Journal of Comparative Psychology*, **108**, 243-251.

Bugnyar, T., Kijne, M., & Kotrschal, K. 2001. Food calling in ravens: are yells referential signals? *Animal Behaviour*, **61**, 949–958.

Charlton, B. D., Reby, D., & McComb, K. 2007a. Female red deer prefer the roars of larger males. *Biology Letters*, **3**, 382-5.

Charlton, B. D., Reby, D., & McComb, K. 2007b. Female perception of size-related formant shifts in red deer, Cervus elaphus. *Animal Behaviour*, **74**, 707–714.

Charrier, I., Mathevon, N., & Jouventin, P. 2003. Vocal signature recognition of mothers by fur seal pups. *Animal Behaviour*, **65**, 543-550.

Cheney, D. L., & Seyfarth, R. M. 1988. Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, **36**, 477–486.

Cheney, D. L., & Seyfarth, R. M. 1990. *How monkeys see the world: inside the mind of another species.* Chicago, IL: University of Chicago Press.

Cheney, D. L., & Seyfarth, R. M. 2007. *Baboon metaphysics: The evolution of a social mind.* Chicago, IL: University of Chicago Press.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., & Maneewongvatana, S. 2008. Encoding emotions in speech with the size code—A perceptual investigation. *Phonetica*, **65**, 210.

Clutton-Brock, T. H., Albon, S., Gibson, R., & Guinness, F. 1979. The logical stag: Adaptive aspects of fighting in red deer (Cervus elaphus L.). *Animal Behaviour*, **27**, 211-225.

Cohen, J. A., & Fox, M. W. 1976. Vocalizations in wild canids and possible effects of domestication. *Behavioural Processes*, **1**, 77-92.

Coppinger, R. P., & Feinstein, M. 1991. “Hark! Hark! the dogs bark.. .” and bark and hark. *Smithonian*, **21**, 119-128.

Coulon, M., Deputte, B. L., Heyman, Y., Delatouce, L., Richard, C., & Baudoin, C. 2007. Visual discrimination by heifers (Bos taurus) of their own species. *Journal of Comparative Psychology*, **121**, 198-204.

Cox, C. R. 1981. Agonistic Encounters Among Male Elephant Seals: Frequency, Context, and the Role of Female Preference. *Integrative and Comparative Biology*, **21**, 197-209.

Crockford, C., & Boesch, C. 2003. Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. *Animal Behaviour*, **66**, 115-125.

Dawkins, R., & Krebs, J. R. 1978. Animal signals: information or manipulation. In: *Behavioural ecology: An evolutionary approach*, pp. 282–309.

Donath, J. 2007. Signals, cues and meaning. In: *Signals, Truth and Design*, MIT Press.

Drahota, A., Costall, A., & Reddy, V. 2008. The vocal communication of different kinds of smile. *Speech Communication*, **50**, 278-287.

Enquist, M., & Leimar, O. 1983. Evolution of fighting behaviour: Decision rules and assessment of relative strength. *Journal of Theoretical Biology*, **102**, 387-410.

Enquist, M., Leimar, O., Ljungberg, T., Mallner, Y., & Segerdahl, N. 1990. A test of the sequential assessment game: fighting in the cichlid fish Nannacara anomala. *Animal Behaviour*, **40**, 1–14.

Evans, C. S., & Evans, L. 1999. Chicken food calls are functionally referential. *Animal behaviour*, **58**, 307–319.

Evans, C. S., & Marler, P. R. 1995. Language and animal communication: Parallels and contrasts. In: *Comparative approaches to cognitive science*, (Ed. by H. L. Roitblat & J. Arcady-Meyer), pp. 341-382. Cambridge: MIT Press.

Evans, C. S., Evans, L., & Marler, P. R. 1993. On the meaning of alarm calls: functional reference in an avian vocal system. *Animal Behaviour*, **46**, 23–38.

Evans, T. A., Howell, S., & Westergaard, G. C. 2005. Auditory--visual cross-modal perception of communicative stimuli in tufted capuchin monkeys (Cebus apella). *Journal of experimental psychology. Animal behavior processes*, **31**, 399-406.

Fagen, R. 1981. *Animal play behaviour.* New York, NY: Oxford University Press.

Fagot, J., Martin-Malivel, J., & Dépy, D. 2000. What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates. In: *Picture perception in animals*, (Ed. by J. Fagot), pp. 295–320. Psychology Press.

Fagot, J., Thompson, R. K. R., & Parron, C. 2010. How to read a picture: lessons from nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 519-20.

Fant, G. 1960. *Acoustic theory of speech production.* Mouton De Gruyter.

Faragó, T., Pongrácz, P., Miklósi, Á., Huber, L., Virányi, Z., & Range, F. 2010a. Dogs' expectation about signalers' body size by virtue of their growls. *PloS one*, **5**, e15175.

Faragó, T., Pongrácz, P., Range, F., Virányi, Z., & Miklósi, Á. 2010b. "The bone is mine": affective and referential aspects of dog growls. *Animal Behaviour*, **79**, 917-925.

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. 2007. Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage*, **36**, 480-7.

Feddersen-Petersen, D. U. 2000. Vocalization of European wolves (Canis lupus lupus L.) and various dog breeds (Canis lupus f. familiaris). *Archive für Tierzucht Dummerstorf*, **43**, 387-397.

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. 2005. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, **69**, 561-568.

Fischer, J. 1998. Barbary macaques categorize shrill barks into two call types. *Animal behaviour*, **55**, 799-807.

Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. 2001a. Acoustic Features of Female Chacma Baboon Barks. *Ethology*, **107**, 33-54.

Fischer, J., Metz, M., Cheney, D. L., & Seyfarth, R. M. 2001b. Baboon responses to graded bark variants. *Animal Behaviour*, **61**, 925–932.

Fitch, W. T. 1994. Vocal tract length perception and the evolution of language. Citeseer.

Fitch, W. T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, **102**, 1213-1222.

Fitch, W. T. 1999. Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *Journal of Zoology*, **248**, 31–48.

Fitch, W. T. 2000. The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. *Phonetica*, **57**, 205-218.

Fitch, W. T., & Hauser, M. D. 2003. Unpacking "Honesty":Vertebrate Vocal Production and the Evolution of Acoustic Signals. In: *Acoustic communication*, (Ed. by A. M. Simmons, A. N. Popper, & R. R. Fay), pp. 65-137. Springer.

Fitch, W. T., & Kelley, J. P. 2001. Perception of vocal tract resonances by whooping cranes Grus americana. *Ethology*, **106**, 559–574.

Fitch, W. T., & Reby, D. 2001. The descended larynx is not uniquely human. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 1669-75.

Fitch, W. T., Neubauer, J., & Herzel, H. 2002. Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, **63**, 407–418.

Flack, J. C., Jeannotte, L. A., & de Waal, F. B. M. 2004. Play signaling and the perception of social rules by juvenile chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, **118**, 149-59.

Fogle, B. 1992. *The dog's mind.* London: Pelham Editions.

Font, E., & Carazo, P. 2010. Animals in translation: why there is meaning (but probably no message) in animal communication. *Animal Behaviour*, **80**, e1-e6.

Fox, M. W. 1970. A comparative study of the development of facial expressions in canids: wolf, coyote and foxes. *Behaviour*, **36**, 49-73.

Fox, M. W. 1971a. *Behaviour of wolves, dogs and related canids.* New York, NY: Harper and Row.

Fox, M. W. 1971b. Socio-infantile and Socio-sexual signals in Canids: a Comparative and Ontogenetic Study. *Zeitschrift für Tierpsychologie*, **28**, 185-210.

Fox, M. W. 1975. Evolution of social behaviour in canids. In: *The wild canids*, (Ed. by M. W. Fox), pp. 429-459. New York, NY: Van Nostrand Reinhold and Company.

Frey, R., & Gebler, A. 2003. The highly specialized vocal tract of the male Mongolian gazelle (Procapra gutturosa Pallas, 1777--Mammalia, Bovidae). *Journal of anatomy*, **203**, 451-71.

Frey, R., Volodin, I. A., & Volodina, E. V. 2007. A nose that roars: anatomical specializations and behavioural features of rutting male saiga. *Journal of anatomy*, **211**, 717-36.

Frey, R., Volodin, I. A., Volodina, E. V., Soldatova, N. V., & Juldaschev, E. T. 2011. Descended and mobile larynx, vocal tract elongation and rutting roars in male goitred gazelles (Gazella subgutturosa Güldenstaedt, 1780). *Journal of anatomy*, **218**, 566-85.

Ghazanfar, A. A., & Logothetis, N. K. 2003. Facial expressions linked to monkey calls. *Nature*, **423**, 937-938.

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D., & Logothetis, N. K. 2007. Vocal-tract resonances as indexical cues in rhesus monkeys. *Current Biology*, **17**, 425-30.

Goodenough, J., McGuire, B., & Jakob, E. Eds. 2001a. Maintaining group cohesion: Description and function of communication and contact. In: *Perspectives on Animal Behavior*, 2nd edn. pp. 421-445. New York, NY: John Wiley & Sons Inc.

Goodenough, J., McGuire, B., & Jakob, E. Eds. 2001b. Maintaining group cohesion: The evolution of communication. In: *Perspectives on Animal Behavior*, 2nd edn. pp. 447-472. New York, NY: John Wiley & Sons Inc.

Greene, E., & Meagher, T. 1998. Red squirrels, Tamiasciurus hudsonicus, produce predator-class specific alarm calls. *Animal behaviour*, **55**, 511–518.

Guihou, A., & Vauclair, J. 2008. Intermodal matching of vision and audition in infancy: A proposal for a new taxonomy. *European Journal of Developmental Psychology*, **5**, 68-91.

Harris, T. R., Fitch, W. T., Goldstein, L. M., & Fashing, P. J. 2006. Black and White Colobus Monkey (Colobus guereza) Roars as a Source of Both Honest and Exaggerated Information About Body Mass. *Ethology*, **112**, 911-920.

Hauser, M. D. 1997. *The evolution of communication.* Cambridge: MIT Press.

Hauser, M. D. 1998. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Animal Behaviour*, **55**, 1647–1658.

Hauser, M. D., & Marler, P. R. 1993. Food-associated calls in rhesus macaques (Macaca mulatta): I. Socioecological factors. *Behavioral Ecology*, **4**, 194–205.

Hauser, M. D., Evans, C. S., & Marler, P. R. 1993. The role of articulation in the production of rhesus monkey (Macaca mulatta) vocalizations. *Animal Behaviour*, **45**, 423–433.

Hawk, S. T., Van Kleef, G. A., Fischer, A. H., & van Der Schalk, J. 2009. "Worth a thousand words": absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, **9**, 293-305.

Hill, H. H. L., & Bekoff, M. 1977. The variability of some motor components of social play and agonistic behaviour in infant eastern coyotes, Canis latrans var. *Animal Behaviour*, **25**, 907-909.

Holekamp, K. E., Boydston, E. E., Szykman, M., Graham, I., Nutt, K. J., Birch, S., Piskiel, A., & Singh, M. 1999. Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. *Animal behaviour*, **58**, 383–395.

Insley, S. J. 2001. Mother-Offspring vocal recognition in northern fur seals is mutual but asymmetrical. *Animal behaviour*, **61**, 129-137.

Janik, V. M., Sayigh, L. S., & Wells, R. S. 2006. Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8293-7.

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. 2005. Monkeys match the number of voices they hear to the number of faces they see. *Current biology : CB*, **15**, 1034-8.

Jouventin, P., & Aubin, T. 2002. Acoustic systems are adapted to breeding ecologies: individual recognition in nesting penguins. *Animal Behaviour*, **64**, 747–757.

Kipper, S., & Todt, D. 2002. The use of vocal signals in the social play of barbary macaques. *Primates; journal of primatology*, **43**, 3-17.

Kiriazis, J., & Slobodchikoff, C. N. 2006. Perceptual specificity in the alarm calls of Gunnison's prairie dogs. *Behavioural Processes*, **73**, 29-35.

Kitchen, D. M., Cheney, D. L., & Seyfarth, R. M. 2003. Female Baboons' Responses to Male Loud Calls. *Ethology*, **109**, 401-412.

Kühme, W. 1963. Ergänzende Beobachtungen an afrikanischen Elefanten (Loxodonta africana Blumenbach 1797) im Freigehege1): (2. Teil). *Zeitschrift für Tierpsychologie*, **20**, 66-79.

Lehner, P. N. 1978. Coyote vocalizations: A lexicon and comparisons with other canids. *Animal Behaviour*, **26**, 712-722.

Lengagne, T., Lauga, J., & Aubin, T. 2001. Intra-syllabic acoustic signatures used by the king penguin in parent-chick recognition: an experimental approach. *The Journal of Experimental Biology*, **204**, 663-672.

Leyhausen, P. 1973. *Verhaltenstudien an Katzen.* Berlin, Germany: Parey.

Lieberman, P. H., & Blumstein, S. 1988. *Speech physiology, speech perception, and acoustic phonetics.* Cambridge: Cambridge University Press.

Liem, K. F. 1985. Ventilation. In: *Function Vertebrate Morphology*, (Ed. by M. Hildebrand), pp. 185–209. Harvard University Press.

Linnankoski, I., Laakso, M.-L., Reijo, A., & Leinonen, L. 1994. Recognition of emotions in macaque vocalizations by children and adults. *Language & Communication*, **14**, 183-192.

Lord, K., Feinstein, M., & Coppinger, R. P. 2009. Barking and mobbing. *Behavioural Processes*, **81**, 358-68.

Macedonia, J. M. 1990. What is communicated in the antipredator calls of lemurs: evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology*, **86**, 177-190.

Macedonia, J. M., & Evans, C. S. 1993. Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology*, **93**, 177–177.

Manser, M. B. 2001. The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2315-24.

Manser, M. B., Seyfarth, R. M., & Cheney, D. L. 2002. Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences*, **6**, 55–57.

Marler, P. R. 1976. Social organization, communication and graded signals: The chimpanzee and the gorilla. In: *Growing points in ethology: based on a conference sponsored by St. John's College and King's College, Cambridge*, (Ed. by P. P. G. Bateson & A. H. Robert), pp. 239. Cambridge: Cambridge University Press.

Maros, K., Pongrácz, P., Bárdos, G., Molnár, C., Faragó, T., & Miklósi, Á. 2008. Dogs can discriminate barks from different situations. *Applied Animal Behaviour Science*, **114**, 159-167.

Martinez, L., & Matsuzawa, T. 2009. Auditory-visual intermodal matching based on individual recognition in a chimpanzee (Pan troglodytes). *Animal cognition*, **12 Suppl 1**, S71-85.

Matsusaka, T. 2004. When does play panting occur during social play in wild chimpanzees? *Primates; journal of primatology*, **45**, 221-9.

Maxwell, C. 1978. Sensitivity and accuracy of the visual analogue scale: a psycho-physical classroom experiment. *British journal of clinical pharmacology*, **6**, 15-24.

Maynard-Smith, J., & Harper, D. 2003. *Animal signals.* Oxford University Press, USA.

McComb, K. 1991. Female choice for high roaring rates in red deer, Cervus elaphus. *Animal Behaviour*, **41**, 79-88.

McElligott, A. G., & Hayden, T. J. 1999. Context-related vocalization rates of fallow bucks, Dama dama. *Animal behaviour*, **58**, 1095–1104.

McElligott, A. G., Birrer, M., & Vannoni, E. 2006. Retraction of the mobile descended larynx during groaning enables fallow bucks (Dama dama) to lower their formant frequencies. *Journal of Zoology*, **270**, 340-345.

Michelsen, A. 1999. The dance language of honeybees: recent findings and problems. In: *The Design of Animal Communication*, (Ed. by M. D. Hauser & M. Konishi),

Mindlin, G. B., & Laje, R. 2005. *The Physics of birdsong.* Springer.

Molnár, C., Kaplan, F., Roy, P., Pachet, F., Pongrácz, P., Dóka, A., & Miklósi, Á. 2008. Classification of dog barks: a machine learning approach. *Animal Cognition*, **11**, 389-400.

Molnár, C., Pongrácz, P., & Miklósi, Á. 2009a. Seeing with ears: Sightless humans' perception of dog bark provides a test for structural rules in vocal communication. *Quarterly journal of experimental psychology (2006)*, 1-10.

Molnár, C., Pongrácz, P., Dóka, A., & Miklósi, Á. 2006. Can humans discriminate between dogs on the base of the acoustic parameters of barks? *Behavioural Processes*, **73**, 76-83.

Molnár, C., Pongrácz, P., Faragó, T., Dóka, A., & Miklósi, Á. 2009b. Dogs discriminate between barks: the effect of context and identity of the caller. *Behavioural Processes*, **82**, 198-201.

Morris, D. 1956. "Typical intensity" and its relation to the problem of ritualisation. *Behaviour*,

Morton, E. S. 1977. On the occurrence and significance of motivation - structural rules in some bird and mammal sounds. *The American Naturalist*, **111**, 855-869.

Mundry, R., & Sommer, C. 2007. Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour*, **74**, 965–976.

Müller, C. A., & Manser, M. B. 2008. Mutual recognition of pups and providers in the cooperatively breeding banded mongoose. *Animal Behaviour*, **75**, 1683-1692.

Narins, P. M., Feng, A. S., Fay, R. R., & Popper, A. N. Eds. 2007. *Hearing and sound communication in amphibians.* Springer Verlag.

Newton-Fisher, N., Harris, S., White, P., & Jones, G. 1993. Structure and function of red fox Vulpes vulpes vocalisations. *Bioacoustics*, **5**, 1–31.

Nicastro, N., & Owren, M. J. 2003. Classification of domestic cat (Felis catus) vocalizations by naive and experienced human listeners. *Journal of Comparative Psychology*, **117**, 44-52.

Notman, H., & Rendall, D. 2005. Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Animal Behaviour*, **70**, 177-190.

Ohala, J. J. 1996. Ethological theory and the expression of emotion in the voice. In: *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, Vol 3 pp. 1812-1815.

Owings, D. H., & Morton, E. S. 1998. *Animal vocal communication: A new approach.* Cambridge University Press.

Owren, M. J., & Rendall, D. 1997. An affect-conditioning model of nonhuman primate vocal signaling. In: *Perspectives in Ethology, Volume 12*, (Ed. by D. H. Ownings, M. D. Beecher, & N. S. Thompson), pp. 299-346. New York: Plenum Publishing Corporation.

Owren, M. J., & Rendall, D. 2003. Salience of caller identity in rhesus monkey (Macaca mulatta) coos and screams: perceptual experiments with human (Homo sapiens) listeners. *Journal of Comparative Psychology*, **117**, 380-90.

Owren, M. J., Seyfarth, R. M., & Cheney, D. L. 1997. The acoustic features of vowel-like grunt calls in chacma baboons (Papio cyncephalus ursinus): Implications for production processes and functions. *The Journal of the Acoustical Society of America*, **101**, 2951-2963.

Palacios, V., Font, E., & Márquez, R. 2007. Iberian wolf howls: acoustic structure, individual variation, and a comparison with north american populations. *Journal of Mammalogy*, **88**, 606-613.

Palagi, E. 2008. Sharing the motivation to play: the use of signals in adult bonobos. *Animal Behaviour*, **75**, 887-896.

Palagi, E. 2009. Adult play fighting and potential role of tail signals in ringtailed lemurs (Lemur catta). *Journal of Comparative Psychology*, **123**, 1-9.

Pellis, S. M., & Pellis, V. C. 1996. On knowing it's only play: The role of play signals in play fighting. *Aggression and Violent Behavior*, **1**, 249-268.

Pellis, S. M., & Pellis, V. C. 2011. To whom the play signal is directed: A study of headshaking in black-handed spider monkeys (Ateles geoffroyi). *Journal of Comparative Psychology*, **125**, 1-10.

Pongrácz, P., Miklósi, Á., Dóka, A., & Csányi, V. 2003. Successful application of video-projected human images for signalling to dogs. *Ethology*, **109**, 809–821.

Pongrácz, P., Molnár, C., & Miklósi, Á. 2006. Acoustic parameters of dog barks carry emotional information for humans. *Applied Animal Behaviour Science*, **100**, 228–240.

Pongrácz, P., Molnár, C., & Miklósi, Á. 2010. Barking in family dogs: an ethological approach. *The Veterinary journal*, **183**, 141-7.

Pongrácz, P., Molnár, C., Miklósi, Á., & Csányi, V. 2005. Human listeners are able to classify dog (Canis familiaris) barks recorded in different situations. *Journal of Comparative Psychology*, **119**, 136-44.

Proops, L., McComb, K., & Reby, D. 2009. Cross-modal individual recognition in domestic horses (Equus caballus). *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 947-51.

Racca, A., Amadei, E., Ligout, S., Guo, K., Meints, K., & Mills, D. S. 2009. Discrimination of human and dog faces and inversion responses in domestic dogs (Canis familiaris). *Animal cognition*,

Rand, A. S., & Dudley, R. 1993. Frogs in helium: the anuran vocal sac is not a cavity resonator. *Physiological Zoology*, **66**, 793-806.

Range, F., Aust, U., Steurer, M., & Huber, L. 2008. Visual categorization of natural stimuli by domestic dogs. *Animal cognition*, **11**, 339-47.

Reby, D., & McComb, K. 2003a. Vocal communication and reproduction in deer. In: *Advances in the Study of Behavior, Volume 33*, Vol 33 (Ed. by H. J. Brockmann, T. J. Roper, M. Naguib, K. E. Wynne-Edwards, J. C. Mitani, & L. W. Simmons), pp. 231–264. Burlington: Academic press.

Reby, D., & McComb, K. 2003b. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, **65**, 519-530.

Reby, D., Hewison, M., Izquierdo, M., & Pépin, D. 2001. Red Deer (Cervus elaphus) Hinds Discriminate Between the Roars of Their Current Harem-Holder Stag and Those of Neighbouring Stags. *Ethology*, **107**, 951-959.

Reby, D., Joachim, J., Lauga, J., Lek, S., & Aulagnier, S. 1998. Individuality in the groans of fallow deer (Dama dama) bucks. *Journal of Zoology*, **245**, 79–84.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C. J., Fitch, W. T., & Clutton-Brock, T. H. 2005. Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 941-7.

Remmers, J. E., & Gautier, H. 1972. Neural and mechanical mechanisms of feline purring. *Respiration Physiology*, **16**, 351-361.

Rendall, D., Rodman, P. S., & Emond, R. E. 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, **51**, 1007–1015.

Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. 1999. The meaning and function of grunt variants in baboons. *Animal Behaviour*, **57**, 583–592.

Riede, T., & Fitch, W. T. 1999. Vocal tract length and acoustics of vocalization in the domestic dog (Canis familiaris). *The Journal of Experimental Biology*, **202**, 2859–2867.

Riede, T., Mitchell, B. R., Tokuda, I. T., & Owren, M. J. 2005. Characterizing noise in nonhuman vocalizations: Acoustic analysis and human perception of barks by coyotes and dogs. *The Journal of the Acoustical Society of America*, **118**, 514.

Riede, T., Tokuda, I. T., Munger, J. B., & Thomson, S. L. 2008. Mammalian laryngeal air sacs add variability to the vocal tract impedance: physical and computational modeling. *The Journal of the Acoustical Society of America*, **124**, 634-47.

Robbins, R. L. 2000. Vocal communication in free-ranging African wild dogs (Lycaon pictus). *Behaviour*, **137**, 1271–1298.

Rogers, L. J., & Andrew, R. J. Eds. 2002. *Comparative vertebrate lateralization.* Cambridge: Cambridge University Press.

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. 2010. Perceptual cues in nonverbal vocal expressions of emotion. *The Quarterly journal of experimental psychology*, **63**, 2251-72.

Schassburger, R. M. 1993. Vocal communication in the timber wolf, Canis lupus, Linnaeus: structure, motivation, and ontogeny. In: *Advances in Ethology*, (Ed. by M. Dehnhard & H. Hofer), Berlin, Germany: Paul Parey Publishers.

Schirmer, A., & Kotz, S. A. 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in cognitive sciences*, **10**, 24-30.

Searby, A., & Jouventin, P. 2003. Mother-lamb acoustic recognition in sheep: a frequency coding. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 1765-71.

Sebeok, T. A. 2007. Coding in the evolution of signalling behavior. *Behavioral Science*, **7**, 430-442.

Seyfarth, R. M., & Cheney, D. L. 1980. The Ontogeny of Vervet Monkey Alarm Calling Behavior: A Preliminary Report. *Zeitschrift für Tierpsychologie*, **54**, 37-56.

Seyfarth, R. M., & Cheney, D. L. 2003. Meaning and Emotion in Animal Vocalizations. *Annals of the New York Academy of Sciences*, **1000**, 32-55.

Seyfarth, R. M., & Cheney, D. L. 2009. Seeing who we hear and hearing who we see. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 669-70.

Seyfarth, R. M., Cheney, D. L., & Marler, P. R. 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, **210**, 801-803.

Shriner, W. M. 1998. Yellow-bellied marmot and golden-mantled ground squirrel responses to heterospecific alarm calls. *Animal behaviour*, **55**, 529–536.

Simmons, A. M., Popper, A. N., & Fay, R. R. Eds. 2003. *Acoustic communication.* Springer.

Siniscalchi, M., Quaranta, A., & Rogers, L. J. 2008. Hemispheric specialization in dogs for processing different acoustic stimuli. *PloS one*, **3**, e3349.

Siniscalchi, M., Sasso, R., Pepe, A. M., Vallortigara, G., & Quaranta, A. 2010. Dogs turn left to emotional stimuli. *Behavioural brain research*,

Slatkin, M., & Smith, J. M. 1979. Models of Coevolution. *The Quarterly Review of Biology*, **54**, 233-263.

Sliwa, J., Duhamel, J.-R., Pascalis, O., & Wirth, S. 2011. Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 1735-1740.

Slocombe, K. E., & Zuberbühler, K. 2005a. Functionally referential communication in a chimpanzee. *Current Biology*, **15**, 1779–1784.

Slocombe, K. E., & Zuberbühler, K. 2005b. Agonistic screams in wild chimpanzees (Pan troglodytes schweinfurthii) vary as a function of social role. *Journal of Comparative Psychology*, **119**, 67-77.

Slocombe, K. E., Kaller, T., Call, J., & Zuberbühler, K. 2010. Chimpanzees Extract Social Information from Agonistic Screams. *PLoS ONE*, **5**, e11473.

Smith, D. R. R., & Patterson, R. D. 2005. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America*, **118**, 3177.

Speaks, C. E. 1992. *Introduction to sound: Acoustics for the hearing and speech sciences.* San Diego, CA: Singular Publishing Group.

Struhsaker, T. T. 1967. Auditory communication among vervet monkeys (Cercopithecus aethiops). In: *Social communication among primates*, (Ed. by S. A. Altmann), Chicago: University of Chicago Press.

Sutton-Smith, B. 2001. *The ambiguity of play.* 2nd edn. Cambridge, MA: Harvard Univ Press.

Számadó, S. 2008. How threat displays work: species-specific fighting techniques, weaponry and proximity risk. *Animal Behaviour*, **76**, 1455-1463.

Számadó, S. 2011. The cost of honesty and the fallacy of the handicap principle. *Animal Behaviour*, **81**, 3-10.

Tamura, N. 1995. Postcopulatory mate guarding by vocalization in the Formosan squirrel. *Behavioral Ecology and Sociobiology*, **36**, 377-386.

Taylor, A. M., & Reby, D. 2010. The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology*, **280**, 221-236.

Taylor, A. M., Reby, D., & McComb, K. 2008. Human listeners attend to size information in domestic dog growls. *The Journal of the Acoustical Society of America*, **123**, 2903-9.

Taylor, A. M., Reby, D., & McComb, K. 2009. Context-Related Variation in the Vocal Growling Behaviour of the Domestic Dog (Canis familiaris). *Ethology*, **115**, 905-915.

Taylor, A. M., Reby, D., & McComb, K. 2010a. Size communication in domestic dog, Canis familiaris, growls. *Animal Behaviour*, **79**, 205-210.

Taylor, A. M., Reby, D., & McComb, K. 2010b. Why Do Large Dogs Sound More Aggressive to Human Listeners: Acoustic Bases of Motivational Misattributions. *Ethology*, **116**, 1155-1162.

Taylor, A. M., Reby, D., & McComb, K. 2011. Cross modal perception of body size in domestic dogs (Canis familiaris). *PloS one*, **6**, e17069.

Tembrock, G. 1976. Canid vocalizations. *Behavioural Processes*, **1**, 57–75.

Titze, I. R. 1994. *Principles of voice production.* Prentice Hall.

Torriani, M. V. G., Vannoni, E., & McElligott, A. G. 2006. Mother-Young Recognition in an Ungulate Hider Species: A Unidirectional Process. *The American naturalist*, **168**, 412–420.

Townsend, S. W., & Manser, M. B. 2011. The function of nonlinear phenomena in meerkat alarm calls. *Biology letters*, **7**, 47-9.

Trivers, R. L. 1971. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, **46**, 35-57.

Truax, B. 2001. *Acoustic communication.* Norwood, New Jersey: Ablex Publishing Corporation.

Vas, J., Topál, J., Gácsi, M., Miklósi, Á., & Csányi, V. 2005. A friend or an enemy? Dogs' reaction to an unfamiliar person showing behavioural cues of threat and friendliness at different times. *Applied Animal Behaviour Science*, **94**, 99-115.

Vettin, J., & Todt, D. 2005. Human laughter, social play, and play vocalizations of non-human primates: an evolutionary approach. *Behaviour*, **142**, 217-240.

Volodin, I. A., Volodina, E. V., & Isaeva, I. V. 2001. Vocal repertoire in the dhole Cuon alpinus (Carnivora, Canidae) in captivity. *Entomological Review*, **81**, S161-S166.

Volodina, E. V., Volodin, I. A., & Filatova, O. A. 2006. The Occurence Of Nonlinear Vocal Phenomena In Frustration Whines Of The Domestic Dog (Canis Familiaris). *Advances in Bioacoustics 2*, **47**, 245-255.

Walker, A. S. 1982. Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, **33**, 514-535.

Wheeler, B. C. 2009. Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings. Biological sciences / The Royal Society*, **276**, 3013-8.

Wheeler, B. C. 2010. Decrease in Alarm Call Response Among Tufted Capuchins in Competitive Feeding Contexts: Possible Evidence for Counterdeception. *International Journal of Primatology*, **31**, 665-675.

Whiten, A., & Byrne, R. W. 1988. Tactical deception in primates. *Behavioral and Brain Sciences*, **11**, 233.

Wilden, I., Herzel, H., Peters, G., & Tembrock, G. 1998. Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics*, **9**, 171–196.

Yeon, S.-C. 2007. The vocal communication of canines. *Journal of Veterinary Behavior: Clinical Applications and Research*, **2**, 141-144.

Yin, S., & McCowan, B. 2004. Barking in domestic dogs: context specificity and individual identification. *Animal Behaviour*, **68**, 343–355.

Yumoto, E., Gould, W. J., & Baer, T. 1982. Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, **71**, 1544–1550.

Zahavi, A., Zahavi, A., Zahavi-Ely, N., & Ely, M. P. 1997. *The handicap principle: A missing piece of Darwin's puzzle.* New York: Oxford University Press, USA.

Zuberbühler, K., Noe, R., & Seyfarth, R. M. 1997. Diana monkey long-distance calls: messages for conspecifics and predators. *Animal Behaviour*, **53**, 589–604.

Summary

This dissertation contains the results of experiments on the communicative aspects of dog growls. This vocalization is a common call type in the Canids, mostly used in agonistic contexts. However, the domestic dogs use it also in play contexts. This functional duality suggests that dog growls possibly carry contextual information.

To test this, we recorded several sequences of growls from dogs in three different contexts: during play, guarding a bone from another dog, and reacting to a threatening human. We analysed the acoustic structure of the growls and found that play growls differ acoustically from the other two agonistic types of growls.

We also performed playback tests in a semi natural food guarding situation. Results showed that food guarding growls deter other dogs more effectively from taking away a seemingly unattended bone than growls recorded in the threatening stranger, or the playful situation, suggesting possible encoded contextual information.

Recently it was found that a dogs' growl contains size cue about the caller. Whether dogs can use this information is as yet unclear. Thus, in the third experiment we tested dogs' ability to assess the size of another dog if they hear a growl paired with projection of two dog pictures. One of them matched the size of the growling dog, while the other one was either 30% larger or smaller. In control groups, noise, cat pictures or projections of triangles were used. The results showed that dogs look sooner and longer at the dog picture matching the size of the caller, when they hear dog growls. No such preference was found with any of the control stimuli, suggesting that dogs have a mental representation of the caller when hearing its vocalization.

Finally, in a playback experiment we tested whether humans can assess the inner state of the growling dogs, and also can they categorize the growls correctly according to their contexts. To achieve this, we played sequences of natural growl samples from our three contexts for 40 adult humans and asked them to rate each growl on emotional scales. We also asked them to guess the possible context of the growls. Our results showed that humans were able to categorize correctly the growls above chance, although they often confused the two agonistic growl types. They rated the growls' emotional background in accordance with the assumed inner states of dogs. Interestingly they found the food guarding growls significantly more aggressive than the growls from the other agonistic context.

In summary, this work provides the first evidence on the possibility of functional referential calls in a large terrestrial carnivore; also we showed first dogs' ability to assess the size of a growling conspecific, and their capability to process crossmodal information. Moreover we showed that humans are able to recognize the emotional background and the contexts of the heard growls based on their acoustical and temporal structure. All these suggest that growls can have an important role in both inter- and intraspecific communication.

Összefoglaló

Disszertációmban a kutya morgásának kommunikációs szerepét vizsgáltuk. Ez a vokalizáció típus általánosan ismert a kutyaféléknél, elsősorban agonisztikus helyzetekben, viszont legalábbis a házi kutyánál jól dokumentált játékos szituációkban is. Ez a funkcionális kettősség arra utal, hogy a morgás hordozhat helyzet specifikus információt.

Ennek vizsgálatához három helyzetben gyűjtöttünk morgásokat: játék közben, amikor egy ember fenyegetően megközelíti a kutyát, és egy csont őrzése közben. A gyűjtött morgások akusztikai paramétereit összehasonlítva, azt találtuk, hogy a játékos morgások jelentősen különböztek a két agonisztikus morgástól. Ezután a morgásokat kutyáknak játszottuk vissza egy táplálék kompetíciós helyzetet modellező kísérletben. Az eredmények szerint a helyzethez illő csontőrző morgások voltak leginkább elrettentő hatással az alanyokra, míg a játékos és a fenyegető idegen helyzetben rögzítettek nem riasztották el a kutyákat a kihelyezett csonttól. Ez megerősíti a feltételezést, hogy a morgások helyzet specifikus információt közölhetnek.

Korábbi vizsgálatok kimutatták, hogy az agresszív morgások utalnak a testméretre, viszont hogy a kutyák képesek-e méretbecslésre használni ezt, az még nem tisztázott. Ezért harmadik kísérletünkben egy morgás bejátszása mellett két kutya fényképet vetítettünk kutyáknak, oly módon hogy az egyik kép a morgó kutyával megegyező méretűnek látszott, míg a másik 30%-al kisebb vagy nagyobb volt. Azt találtuk, hogy a kutyák hamarabb és hosszabban néztek a méretben illeszkedő képre, viszont ez a preferencia eltűnt, ha a morgás helyett zajt, vagy ha a nem kutya képet, hanem háromszögeket vagy macska képeket mutattunk. Ez arra utal, hogy a kutyák a morgást hallva fel tudták mérni a morgó kutya méretét.

Végül arra is kíváncsiak voltunk, hogy az emberek képesek-e érzelmet kapcsolni a morgásokhoz, és a kontextusuk alapján kategorizálni azokat. Ehhez a három helyzetben rögzített morgásokat visszajátszottuk 40 felnőttnek, hogy skálázzák mennyire jellemzőek rájuk egyes érzelmek, és tippeljék meg milyen helyzetből származhattak. A résztvevőknek sikerült a véletlen szint fölött eltalálni a kontextusokat, bár a két agresszív helyzetet gyakran felcserélték. Emellett a helyzetnek megfelelő érzelmi hátteret tulajdonították a morgásoknak, emellett a csontőrző morgást az emberek jóval haragosabbnak értékelték, mint a fenyegető idegenre morgást.

Összefoglalva: a disszertációmban taglalt vizsgálatok elsőként utalnak funkcionális referencialitás lehetőségére egy nagytestű szárazföldi ragadozó kommunikációs

rendszerében, valamint arra, hogy a kutyák képesek pontosan felmérni egy fajtárs méretét pusztán a morgása alapján, és képesek modalitások közt ezt az információt összekapcsolni. Végezetül azt is kimutattuk, hogy az emberek képesek megfelelő érzelmi hátteret tulajdonítani a különböző helyzetből származó morgásoknak, és a helyzetük alapján képesek azokat a kategorizálni is. Mindezek arra utalnak, hogy a morgások fontos szerepet játszhatnak mind az inter-, mind az intraspecifikus kommunikációban.

Appendix

Table 5. The parameters of the dogs used for sound recording and the contexts of the successful recordings (Experiment 1).

Name	Breed	Sex	Age (Year)	Weight (Kg)	Height At Withers (Cm)	Type Of Growl Recorded
Angel	Mudi	female	7.0	13	44	FG, TS
Barka	Border collie	male	2.5	19	56	TS,PL
Bodza	Pumi	female	2	18	48	PL
Boskó	Magyar vizsla	male	10.0	27	63	FG, PL
Chili	Mudi	female	2.5	15	46	TS
Cooper	Border collie	male	5.0	20	54	TS
Edgar	Foxterrier	male	4.0	7.5	40	FG
Fecske	Mudi	female	7.0	17	50	FG, TS, PL
Gréti	Schnauzer	female	3.0	17	43	PL
Guru	Belgian sheperd	male	2.5	29	64	FG
Jamile	Mongrel	male	4.0	35	60	FG, TS, PL
Jenny	Mongrel	female	3.0	25	55	TS, PL
Jóság	Mudi	female	0.8	10	46	TS
Kevin	Mongrel	male	5.5	32	61	FG, TS, PL
Kicsifeka	Mongrel	male	3.0	18	30	TS
Kira	Mongrel	female	4.5	12	45	FG
Lili	Magyar vizsla	female		19	59	TS, PL
Linka	Mongrel	female	5.0	34	64	FG, TS
Lucy	Border collie	female	4.0	18	52	FG, TS, PL
Mio	Mongrel	male	2.5	30	60	FG, TS, PL
Tódor	Mongrel	male	6.0			FG
Totó	West highland terrier	male	4.0	10	35	TS, PL
Vacak	Pumi	male	1.5	15	46	TS
Xena	Jackrussel terrier	female	6.0	8.5	32	FG

Table 6. The descriptives of the acoustical parameters of the growls, and the statistical results: Mixed Effects Model with Tukey post hoc test (Experiment 1).

	FG		TS		PL		MEM		Post hoc test		
Acoustical parameters	***mean***	***SD***	***mean***	***SD***	***mean***	***SD***	$F_{2;18}$	***P***	**FG vs. TS**	**FG vs. PL**	**TS vs. PL**
Length (s)	1.53	0.21	1.36	0.18	0.68	0.09	**4.103**	**0.04**	0.782	**0.008**	**0.021**
Fundamental frequency (Hz)	178.4	51.13	150	49.69	506.1	114.1	**4.355**	**0.034**	0.96	**0.017**	**0.004**
Formant1 (Hz)	631.8	61.3	561.2	33.38	886.1	51.02					
Formant2 (Hz)	1934	100.2	1814	52.74	1838	53.98					
Formant3 (Hz)	3236	71.06	3102	77.46	2971	57.56					
Formant4 (Hz)	4526	47.74	4403	56.23	4258	81.41					
Formant5 (Hz)	5380	18.03	5347	15.71	5304	23.75					
Formant dispersion (Hz)	1187	15.87	1197	8.4	1104	17.24	**9.638**	**0.002**	0.871	**0.001**	**>0.0001**
Standard Deviation of Formants (Hz)	334	25.05	314.2	25.28	292	33.42	0.339	0.718			
Harmonic to Noise Ratio	3.64	0.88	2.82	0.63	1.88	0.54	1.489	0.259			

Table 7. The parameters of the subjects and the type of heard growl (Experiment 2).

Name	Breed	Sex	Age (Year)	Weight (Kg)	Height At Withers (Cm)	Type Of Growl Heard
Angel	Mudi	female	7.0	13	44	linkaFG
Barnabás	Mongrel	male	1.0	27	60	fecskeTS
Bastian	Border collie	male	2.0	25	55	mioFG
Benji	Collie	male	5.0	23	63	barkaPL
Benji	Mongrel	male	10.0	30	66	bodzaPL
Berci	Mongrel	male	4.0	31	67	bodzaPL
Bertl	Australian sheperd	male	5.0	26	60	mioPL
Boris	Mongrel	male	6.0	30	62	cooperTS
Chilli	Mudi	female	2.5	15	46	linkaTS
Cooper	Border collie	male	5.0	20	54	jamilPL
Cora	Staffordshire terrier	female	2.5	26	53	fecskeFG
Diron	Belgian sheperd	male	2.0	24	62	linkaFG
Füli	Mongrel	male	5.0	10	42	fecskeTS
Grimbusz	Belgian sheperd	male	1.5	28	64	fecskeFG
Guiness	Border collie	female	3.5	17	41	jamilFG
Guru	Belgian sheperd	male	2.5	29	64	mioTS
Hanga	Mongrel	female	1.5	20	59	fecskePL
Jamile	Mongrel	male	4.0	35	60	mioFG
Jana	Mongrel	female	2.0	22	62	cooperTS
Jeffi	Mongrel	female	11.5	21	53	fecskePL
Jersey	Border collie	female	1.5	14	47	linkaTS
Kira	Mongrel	female	4.5	12	45	jennyPL
Kutyácska	Mongrel	female	12.0	12	23	guruFG
Kyra	Mongrel	female	12.0	24	55	jamilTS
Lidi	Mongrel	female	10.0	23	55	angelTS
Lili	Collie	female	6.0	21	57	barkaPL
Lujza	Mongrel	female	11.0	35	45	angelTS
Morzsi	Mongrel	female	11.0	12	42	guruFG
Rainbow	Border collie	male	1.0	14	46	kevinFG
Sue	Mongrel	female	6.0	15	50	mioTS
Summer	Border collie	female	3.0	16	45	jamilTS
Swennie	Border collie	female	3.5	17	41	kevinFG
Szöszke	Poodle	female	1.0	6,5	43	jennyPL
Tódor	Mongrel	male	5	15	37	jamilFG
Vito	Australian sheperd	male	1	29	56	mioPL
Zsivány	Magyar vizsla	male	4.0	28	61	jamilPL

Table 8. The group arrangement of the subjects, and the parameters of the presented growls and pictures (Experiment 3).

group	name of subject	Breed of subject	sex of subject	growl ID	sex of the growling dog	size of the subject	Size category of the growling dog	Size of the growling dog in cm	Size of the modified picture in cm	Sizing the modified picture	side of the modified picture
Dog-Noise	Toffee	Golden retriever	female	4	female	Large	Small	50	35	Smaller	Left
	Guiness	Border Collie	female	1	female	Small	Small	44	30.8	Smaller	Left
	Yukie	Labrador retriever	male	9	female	Large	Small	52	36.4	Smaller	Left
	Candy	Labrador retriever	female	10	female	Large	Large	60	42	Smaller	Right
	Luna	Border Collie	female	5	male	Small	Large	63	44.1	Smaller	Right
	Ida	German shepherd	female	11	male	Large	Large	61	42.7	Smaller	Right
	Yogibare	Yorkshire terrier	male	2	male	Small	Small	45	31.5	Smaller	Right
	Penelope	Labrador retriever	female	3	female	Large	Small	32	22.4	Smaller	Right
	Lisa	Golden retriever	female	8	male	Large	Small	40	28	Smaller	Right
	Tiny	Foxterrier	male	6	female	Small	Large	64	44.8	Smaller	Left
	Naib	Österreicher pintscher	male	7	male	Small	Large	64	44.8	Smaller	Left
	Feder	Border Collie	female	12	male	Large	Large	60	42	Smaller	Left
	Tinkerbell	Zwergpintscher	female	10	male	Small	Large	60	78	Bigger	Right
	Jacks	Mongrel	male	5	male	Small	Large	63	81.9	Bigger	Right
	Lucky	Mongrel	male	11	male	Small	Large	61	79.3	Bigger	Right
	Flamme	Mongrel	female	2	female	Large	Small	45	58.5	Bigger	Left
	Lilly	Zwergpintscher	female	3	female	Small	Small	32	41.6	Bigger	Left

	Cini	Zwergpintscher	female	8	male	Small	Small	40	52	Bigger	Left
	Minnie	Chinese crestdog	female	4	female	Small	Small	50	65	Bigger	Right
	Palmira	Mongrel	female	1	female	Small	Small	44	57.2	Bigger	Right
	Finn2	Border Collie	male	9	female	Large	Small	52	67.6	Bigger	Right
	Edgar	Foxterrier	male	6	female	Small	Large	64	83.2	Bigger	Left
	Akina	Akita Inu	female	7	male	Large	Large	64	83.2	Bigger	Left
	Pedro	Mongrel	male	12	male	Small	Large	60	78	Bigger	Left
Dog-Growl	Sayon	Australian Shepherd	male	2	female	Large	Small	45	31.5	Smaller	Left
	Tódor	Mongrel	male	4	female	Small	Small	50	35	Smaller	Left
	Achia	Kleiner Münsterländer	female	1	female	Small	Small	44	30.8	Smaller	Left
	Csillag	Kuvasz	female	10	male	Large	Large	60	42	Smaller	Left
	Sissi	Mongrel	female	11	male	Small	Large	61	42.7	Smaller	Left
	Amira	Rhodesian Ridgeback	female	5	male	Large	Large	63	44.1	Smaller	Left
	Happy	Zwergpintscher	female	9	female	Small	Small	52	36.4	Smaller	Right
	Lucy3	Border Collie	female	3	female	Small	Small	32	22.4	Smaller	Right
	Lestat	German shepherd	male	8	male	Large	Small	40	28	Smaller	Right
	Benji	White Swiss Shepherd Dog	male	6	female	Large	Large	64	44.8	Smaller	Right
	Nash	Siberian husky	male	7	male	Large	Large	64	44.8	Smaller	Right
	Nessy	Mongrel	female	12	male	Small	Large	60	42	Smaller	Right
	Lucy	Mongrel	female	2	female	Large	Small	45	58.5	Bigger	Left
	Ted	Mongrel	male	3	female	Small	Small	32	41.6	Bigger	Left

	Havanna	Beagle	female	8	male	Small	Small	40	52	Bigger	Left
	Benji2	Border Collie	male	6	female	Small	Large	64	83.2	Bigger	Left
	Jayjay	Border Collie	female	7	male	Small	Large	64	83.2	Bigger	Left
	Lucy2	Rotweiler	female	12	male	Large	Large	60	78	Bigger	Left
	Jocó	Hungarian vizsla	male	4	female	Large	Small	50	65	Bigger	Right
	Cash	Australian Shepherd	male	1	female	Large	Small	44	57.2	Bigger	Right
	Hugo	Mongrel	male	9	female	Small	Small	52	67.6	Bigger	Right
	Quent	Golden retriever	male	10	male	Large	Large	60	78	Bigger	Right
	Quentin	Border Collie	male	5	male	Large	Large	63	81.9	Bigger	Right
	Chilly	Jackrussel terrier	male	11	male	Small	Large	61	79.3	Bigger	Right
Shape-Growl	Fipsy	Mongrel	male	2	female	Small	Small	45	31.5	Smaller	Right
	Kimmy	Foxterrier	female	3	female	Small	Small	32	22.4	Smaller	Right
	Bolita	Eurasian	female	8	male	Large	Small	40	28	Smaller	Right
	Crissy	Australian Shepherd	female	6	female	Large	Large	64	44.8	Smaller	Left
	Finn	Australian Shepherd	male	7	male	Large	Large	64	44.8	Smaller	Left
	Bounty	Australian Shepherd	female	12	male	Large	Large	60	42	Smaller	Left
	Bony	Belgian shepherd	female	4	female	Large	Small	50	35	Smaller	Left
	Kendra	Belgian shepherd	female	1	female	Large	Small	44	30.8	Smaller	Left
	Oszkár	Mongrel	male	9	female	Small	Small	52	36.4	Smaller	Left
	Amy	Border Collie	female	10	male	Small	Large	60	42	Smaller	Right
	Elaine	Border Collie	female	5	male	Small	Large	63	44.1	Smaller	Right

	Diana	White Swiss Shepherd Dog	female	11	male	Large	Large	61	42.7	Smaller	Right
	Idefix	West Highland Terrier	male	2	female	Small	Small	45	58.5	Bigger	Left
	Izzy	Chihuahua	female	3	female	Small	Small	32	41.6	Bigger	Right
	Alia	Rhodesian Ridgeback	female	8	male	Large	Small	40	52	Bigger	Left
	Shaggy	Mongrel	male	6	female	Small	Large	64	83.2	Bigger	Right
	Beluga	Labrador retriever	female	7	male	Large	Large	64	83.2	Bigger	Left
	Susanne	Mongrel	female	12	male	Large	Large	60	78	Bigger	Right
	Bluebell	Border Collie	female	4	female	Small	Small	50	65	Bigger	Right
	Bobby	French bulldog	male	1	female	Small	Small	44	57.2	Bigger	Left
	Benno	Golden retriever	male	9	female	Large	Small	52	67.6	Bigger	Right
	Nicky	Mongrel	male	10	male	Small	Large	60	78	Bigger	Left
	Kyra	Giant poodle	female	5	male	Large	Large	63	81.9	Bigger	Right
	Lynette	Giant poodle	female	11	male	Large	Large	61	79.3	Bigger	Left
Cat-Growl	Allegro	Border Collie	male	4	female	Large	Small	50	35	Smaller	Right
	Flocke	Border Collie	female	1	female	Small	Small	44	30.8	Smaller	Right
	Benni	Mongrel	male	9	female	Large	Small	52	36.4	Smaller	Right
	Filon	Australian Shepherd	male	2	female	Large	Small	45	31.5	Smaller	Left
	Jimmy	Dalmatiner	male	3	female	Small	Small	32	22.4	Smaller	Left
	Herbie	Mongrel	male	8	male	Large	Small	40	28	Smaller	Left
	Sari	Mongrel	female	4	female	Small	Small	50	35	Smaller	Left
	Ike	Mongrel	male	1	female	Small	Small	44	30.8	Smaller	Left
	Sidney	Border Collie	male	9	female	Large	Small	52	36.4	Smaller	Left

	Sirius	Newfoundlander	male	2	female	Large	Small	45	31.5	Smaller	Right
	Chilly2	Australian Shepherd	female	3	female	Large	Small	32	22.4	Smaller	Right
	Ginger	Mongrel	female	8	male	Large	Small	40	28	Smaller	Right
	Cooper	Mongrel	male	2	female	Large	Small	45	58.5	Bigger	Right
	Foxy	Mongrel	female	3	female	Small	Small	32	41.6	Bigger	Right
	Chico	Shi tzu	male	8	male	Small	Small	40	52	Bigger	Right
	Flyte	Border Collie	male	4	female	Small	Small	50	65	Bigger	Left
	Cooky	Mongrel	female	1	female	Small	Small	44	57.2	Bigger	Left
	Bertl	Australian Shepherd	male	9	female	Large	Small	52	67.6	Bigger	Left
	Kalle	Belgian shepherd	male	2	female	Large	Small	45	58.5	Bigger	Left
	Bandita	Mongrel	male	3	female	Small	Small	32	41.6	Bigger	Left
	Gumó	Mongrel	female	8	female	Small	Small	40	52	Bigger	Left
	Marty	Mongrel	male	4	female	Large	Small	50	65	Bigger	Right
	Xena	Jackrussel terrier	female	1	female	Small	Small	44	57.2	Bigger	Right
	Lottie	Border Collie	female	9	male	Small	Small	52	67.6	Bigger	Right

Table 9. The background variables of the subjects and the ID of the heard growls sessions (Experiment 4)

subject id	gender	age	have dog	bitten	session1	session2
1	2	23	1	0	1	20
2	2	28	1	0	2	19
3	2	28	1	0	3	18
4	2	23	0	0	4	17
5	1	21	1	1	5	16
6	2	33	1	1	6	15
7	2	21	1	1	7	14
8	2	24	1	0	8	13
9	1	54	1	0	9	12
10	2	53	1	1	10	11
11	2	23	1	1	11	10
12	2	20	0	1	12	9
13	2	19	0	0	13	8
14	1	20	1	0	14	7
15	2	29	1	1	15	6
16	1	26	1	1	16	5
17	2	26	0	0	17	4
18	1	26	1	0	18	3
19	2	22	1	1	19	2
20	1	21	1	0	20	1
21	1	26	1	1	20	1
22	2	26	1	1	19	2
23	1	28	1	0	18	3
24	1	29	1	0	17	4
25	1	26	1	1	16	5
26	1	26	0	0	15	6
27	1	27	0	0	14	7
28	1	22	1	0	13	8
29	2	22	1	0	12	9
30	2	27	1	0	11	10
31	2	24	1	1	10	11
32	2	23	0	0	9	12
33	1	24	0	0	8	13
34	2	21	0	0	7	14
35	2	30	1	1	6	15
36	2	20	1	0	5	16
37	2	22	1	0	4	17
38	2	23	1	1	3	18
39	2	37	1	0	2	19
40	2	20	1	1	1	20

Table 10. The gender and the size parameters of the growling dogs and the acoustical parameters (F0: fundamental frequency; dF: formant dispersion) of the used growl samples in the playback (Experiment 4).

name	context	growl id	sex	weight	height	growl N	growl length (s)	F0 (Hz)	dF (Hz)
Angel	FG1	a1	2	13	44	7	0.73	67.00	583.25
Boskó	FG2	a2	1	27	63	6	1.37	79.20	881.00
Fecske	FG3	a3	2	17	50	4	1.48	101.00	444.75
Guru	FG4	a4	1	29	64	4	2.10	120.75	1112.67
Kevin	FG5	a5	1	32	61	4	1.71	75.25	754.33
Linka	FG6	a6	2	34	64	4	2.26	66.00	938.50
Mio	FG7	a7	1	30	60	4	2.03	99.00	499.67
Kira	FG8	a8	2	12	45	3	3.26	125.00	647.25
Angel	TS1	b1	2	13	44	7	0.94	74.83	585.00
Chili	TS2	b2	2	15	46	2	4.30	61.00	740.75
Cooper	TS3	b3	1	20	54	4	1.96	86.25	687.00
Fecske	TS4	b4	2	17	50	5	1.43	67.31	928.25
Jóság	TS5	b5	2	10	46	3	2.48	102.10	832.25
Kevin	TS6	b6	1	32	61	6	1.16	117.02	1022.00
Linka	TS7	b7	2	34	64	3	3.11	66.89	545.00
Toto	TS8	b8	1	10	35	6	1.01	120.00	911.25
Barka	PL1	c1	1	19	56	9	0.95	117.00	797.33
Boskó	PL2	c2	1	27	63	6	1.44	104.74	451.67
Bodza	PL3	c3	2	18	48	20	0.39	181.20	448.33
Gréti	PL4	c4	2	17	43	6	0.75	91.88	636.00
Lili	PL5	c5	2	19	59	16	0.47	108.70	864.00
Fecske	PL6	c6	2	17	50	12	0.62	97.10	651.50
Jamile	PL7	c7	1	35	60	16	0.42	112.10	545.00
Jenny	PL8	c8	2	25	55	14	0.65	129.70	477.00

Table 11. The order of growl samples in the 20 playback sessions (Experiment 4).

Session	growl ID						
	1	**2**	**3**	**4**	**5**	**6**	***7***
1	b6	c8	c7	a5	b2	a4	*b6*
2	c1	a6	b6	c6	b7	a4	*c1*
3	a5	c7	c3	b4	b8	a6	*a5*
4	c5	b4	c3	a7	a8	b3	*c5*
5	a6	c5	b8	b6	a1	c1	*a6*
6	b5	c2	c6	a5	a4	b4	*b5*
7	c4	a7	c2	a6	b4	b7	*c4*
8	a4	c8	c4	b6	a5	b7	*a4*
9	a8	c1	a6	c2	b6	b4	*a8*
10	a2	c3	a5	b7	c8	b5	*a2*
11	a7	c5	c1	a3	b2	b1	*a7*
12	c8	b3	c4	b1	a7	a1	*c8*
13	b7	a8	b2	a1	c7	c5	*b7*
14	c3	c8	a7	a1	b2	b5	*c3*
15	b1	b8	c2	c6	a1	a2	*b1*
16	a3	c1	c6	b5	b2	a2	*a3*
17	c7	c5	b1	b3	a8	a3	*c7*
18	b8	a2	b5	a8	c4	c7	*b8*
19	b3	c6	b1	a2	c4	a3	*b3*
20	c2	a3	b8	a4	c3	b3	*c2*